

Université Paris Descartes
Ecole doctorale Génétique Cellule Immunologie Infectiologie Développement
Faculté de Médecine
Et
Université de Montréal
Département de Physiologie, Programme des sciences neurologiques
Faculté de Médecine

Thèse présentée à l'Université Paris Descartes
en vue de l'obtention du grade de Docteur es Sciences
Spécialité Génétique
ET
à la Faculté des études supérieures de l'Université de Montréal
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en sciences neurologiques.

Etude des déterminants génétiques des psychoses à début précoce
Génétique de la schizophrénie et hypothèse glutamatergique

Par

TARABEUX Julien

Soutenue le 12 juillet 2011

Jury

Mr Thomas BOURGERON, Université Paris Diderot	Président
Mme Frédérique BONNET-BRILHAUT, Université Tours I	Rapporteur
Mr Dominique CAMPION, Université Rouen	Rapporteur
Mme Caroline DEMILY, Université Lyon I	Examineur
Mme Marie-Odile KREBS, Université Paris Descartes	Directeur de thèse
Mr Guy A. ROULEAU, Université de Montréal	Directeur de thèse

© TARABEUX Julien, 2011

Table des matières

Remerciements	5
Avant-propos	6
Chapitre 1 : Rappels bibliographiques.....	7
Première partie : La schizophrénie	8
I-Les symptômes	8
II-L'hétérogénéité d'un patient à l'autre, d'un moment à l'autre	9
III-L'hypothèse « neurodéveloppementale »	10
A-Les études familiales	12
B-Les études de jumeaux	14
C-Les études d'adoption.....	15
D-Paradoxe évolutionniste	15
Deuxième partie : Les facteurs génétiques dans la schizophrénie	17
I - Définitions	17
II- La recherche des marqueurs moléculaires	22
A-Études de liaison	22
B-Études d'association, approche gènes candidats	22
C-Études d'association pangénomique	23
D-Étude des anomalies chromosomiques et les variants du nombre de copie	26
E-Recherche des mutations rares	31
F-Problèmes nosographiques : vers un continuum schizophrénie, troubles bipolaires, troubles du spectre autistique	31
G- Approche endophénotypique :	33
Troisième partie : Glutamate et schizophrénie	36
I-Systèmes glutamatergiques	36
A-Introduction	36
B-Voies de synthèse du glutamate	37
C-Les récepteurs glutamatergiques	38
D- Synapse glutamatergique et plasticité synaptique	39
II- Récepteurs NMDA : Description et propriétés	42
III-NMDA et schizophrénie	47
A-Introduction à l'hypothèse glutamatergique	47
B-Arguments pharmacologiques	48
C-Arguments physiopathologiques	49
D-Arguments génétiques.....	50
Chapitre 2 : travail expérimental.....	54
Contexte : le projet « synapse to disease » : S2D	55
Article 1 : De Novo Truncating Mutation in KIF17 Associated with Schizophrenia.	58
Article 2 : Rare mutations in N-méthyl-D-aspartate (NMDA) glutamate receptors in autism spectrum disorders and schizophrenia.	70
Chapitre 3 : Discussion générale.....	106
Annexes	118
Article 3 : Direct measure of the de novo mutation rate in autism and schizophrenia cohort.....	119

Article 4 : Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia.	129
<i>Bibliographie</i>	144

Table des illustrations

Figure 1 : Hérité et risque de schizophrénie.....	13
Figure 2 : Illustration des résultats d'une étude de liaison.....	18
Figure 3 : Illustration du principe des études d'association pangénomique.	20
Figure 4 : Représentation de la diversité génétique générée par les CNV.....	21
Figure 5 : Résumé des études d'association pangénomique de la schizophrénie	26
Figure 6 : Voies de synthèse du glutamate	37
Figure 7 : Récepteurs ionotropiques au glutamate.....	38
Figure 8 : La synapse glutamatergique	39
Figure 9 : Récepteurs glutamatergiques et plasticité synaptique	40
Figure 10 : Structure des récepteurs NMDA	44
Figure 11 : Systèmes GABA et glutamatergiques.	46
Figure 12 : Etudes d'association et sous-unités du récepteur NMDA.....	50
Figure 13 : Synapse glutamatergique et schizophrénie.....	52
Figure 14 : présentation du projet S2D	57

Remerciements

Je voudrais tout d'abord remercier *Marie-Odile Krebs*, *Thérèse Jay* et *Guy Rouleau* de m'avoir accueilli au sein de leurs laboratoires et pour leur aide.

Je voudrais aussi remercier *l'Université Paris Descartes*, *l'Université de Montréal* et le groupe de recherche universitaire du médicament de l'Université de Montréal (*GRUM*) pour m'avoir accordé un financement pour le déroulement de cette thèse.

Un grand merci à *Julie Gauthier* pour m'avoir encadré lors de mes « débuts » à Montréal, à *Fadi Hamdan*, *Amélie Piton*, *Hussein Daoud*, *Ron Lafrenière* pour leurs conseils scientifiques précieux.

A l'équipe de *Pierre Drapeau*, et en particulier à *Nathalie Champagne* sans lesquels la validation fonctionnelle sur le zebrafish n'aurait pu être effectuée.

A ceux qui m'ont accompagné durant deux ans formidables à Montréal en particulier *Hussein* et *Isabelle*, *Hélène* et *Alex* mais aussi tous les membres du laboratoire, *Edouard*, *Dan*, *Ousmane*, *Daniel*, *Pascale*, *Mike*, *Raphael*, *Sylvia*, les deux *Annie*, les deux *Caroline*, *Claude*.

A *Oussama Kebir* qui m'a beaucoup aidé et soutenu ces derniers mois.

A tous les membres du laboratoire et du CERC en France : *Justine*, *Dorian*, *Gwen*, *Sophie*, *Julie*, *Danlei*, *Elisabeth*, *Julie*, *Arnaud*, *Caroline*, *Bill* et à *Marion* en particulier.

A mes parents.

A tous un grand merci

Avant-propos

Ce travail a pour but de fournir une revue non exhaustive de l'ensemble des données existantes concernant la génétique de la schizophrénie afin de fournir une réflexion globale, à la vue des découvertes récentes, sur l'évolution des hypothèses avancées, à la fois génétiques et physiopathologiques, et de tenter de fournir des pistes à l'étiologie de cette maladie qui touche jusqu'à 1% de la population à travers le monde, avec des conséquences lourdes pour les individus atteints, leurs familles et le système de santé.

Après une brève présentation de la schizophrénie et une revue des études génétiques concernant cette maladie, je détaillerai dans ce travail les différents arguments en faveur d'une implication glutamatergique et en particulier des récepteurs N-méthyl-D-aspartate (NMDA) dans l'étiologie de la schizophrénie. Je présenterai ensuite les travaux auxquels j'ai contribué au cours de ma thèse. Enfin, dans une dernière partie, j'essaierai de fournir une réflexion globale sur l'implication des facteurs génétiques dans l'étiologie de cette maladie.

Chapitre 1 : Rappels bibliographiques

Première partie : La schizophrénie

La schizophrénie est une maladie complexe, apparaissant à l'adolescence, causée à la fois par des facteurs génétiques, environnementaux et par leur interaction. La schizophrénie a été considérée comme une maladie distincte depuis plus d'un siècle mais, en l'absence de marqueurs biologiques certains, le diagnostic est clinique, posé sur la constatation de la présence des signes cliniques ou symptômes.

I-Les symptômes

La schizophrénie est une maladie caractérisée par une grande diversité de symptômes que l'on peut répertorier selon trois grandes classes.

Les symptômes positifs regroupent un ensemble d'anomalies psychiques et comportementales qui se manifestent par une exacerbation des fonctions « normales » d'un individu. Cela inclut de façon non exhaustive : illusions, hallucinations, idées délirantes, comportement désorganisé ou inapproprié.

Les symptômes négatifs, au contraire, reflètent le déclin de fonctions « normales » et regroupent des déficits de communication et des troubles du langage (alogie), un retrait social et globalement un émoussement affectif, une difficulté à prendre du plaisir (anhédonie) et un manque de motivation à effectuer une tâche nouvelle ou à effectuer un comportement organisé (avolition).

Enfin, la désorganisation psychique, troubles du cours de la pensée, relâchement des associations, dissociation ou discordance idéo affective.

Ces symptômes sont presque toujours associés à des troubles cognitifs qui concernent, en particulier, des déficits de la mémoire de travail, des déficits de la mémoire visuelle à court terme, des déficits de la mémoire sémantique, des déficits dans la vitesse des processus psychiques, des déficits de raisonnement, des troubles de la vigilance et de l'attention ainsi que des déficits de filtrage des informations sensori-motrices.

Les patients peuvent également présenter des troubles de l'humeur (symptômes dépressifs ou excitation maniaque), de l'anxiété ou des manifestations anxieuses (symptômes pseudo obsessionnels), de l'hostilité, des symptômes catatoniques (moteurs, négativisme, opposition).

II-L'hétérogénéité d'un patient à l'autre, d'un moment à l'autre

La schizophrénie présente une grande hétérogénéité interindividuelle et au cours de l'évolution.

L'âge de début de la maladie est variable selon les individus avec un pic de fréquence d'apparition pouvant grossièrement être placé entre 15 et 30 ans. De plus, la prépondérance des différents symptômes est elle aussi variable au cours du temps et selon les individus, et les patients répondent de façon très inégale au traitement.

Selon les critères du DSM-IV-TR, il existe cinq types majeurs de schizophrénie, le diagnostic reposant sur le type de symptômes que présente la personne au moment de l'évaluation. On parle de schizophrénie de type paranoïde, désorganisée, catatonique, indifférencié et résiduel. Il existe aussi un certain nombre de formes dites « limites » tels que les troubles schizo-affectifs. Cependant, même dans ces sous-groupes, en raison de

l'hétérogénéité des symptômes, les patients ne forment pas de groupes totalement homogènes et les catégories diagnostiques, même si elles visent à caractériser des groupes d'individus présumés homogènes, ont été établies afin de fournir un cadre de travail reproductible pour les études plutôt que de définir des catégories basées sur des processus physiopathologiques identifiés et spécifiques.

III-L'hypothèse « neurodéveloppementale »

Une des principales hypothèses proposée pour expliquer l'étiologie de la maladie est l'idée de perturbations du développement cérébral normal. La supposition principale est que le développement cérébral est atteint de manière spécifique, à des périodes critiques du développement et que les lésions résultantes provoquent les symptômes de la maladie seulement par le biais des processus normaux de maturation cérébrale ayant lieu chez l'adolescent ou le jeune adulte.

Pour simplifier, l'hypothèse « neurodéveloppementale » postule que les patients ayant un diagnostic de schizophrénie à l'âge adulte ont connu une perturbation dans le développement ordonné du cerveau des années avant la phase symptomatique de la maladie, en particulier durant les phases précoces du développement cérébral. Par exemple, les sujets à haut risque génétique (enfants de patients par exemple) présentent des anomalies physiques et neurologiques mineures ainsi que des altérations mineures du fonctionnement cognitif et du développement social, bien avant l'éclosion des symptômes, reflétant des perturbations du développement « normal » (Krebs & Mouchet 2007). De même, les évaluations faites dans l'enfance de sujets développant

ultérieurement des troubles psychotiques montrent souvent des anomalies de la maturation neurologique, des troubles de l'apprentissage voire des troubles psychiatriques survenus dans l'enfance et n'ayant pas donné lieu à un repérage psychopathologique adapté (Rapoport et al 2005). Certaines formes de schizophrénie pourraient ainsi être plus « neurodéveloppementales » que d'autres et ces patients pourraient se situer sur un continuum entre troubles schizophréniques et troubles du spectre autistique (Craddock & Owen 2010). Les schizophrénies à début précoce ou très précoce en particulier pourraient constituer clairement des formes « neurodéveloppementales », ces patients présentant des formes pathologiques à début souvent insidieux, des anomalies prémorbides du neurodéveloppement, des antécédents familiaux plus fréquents, une mauvaise réponse au traitement et un plus mauvais pronostic (Margari et al 2011; Margari et al 2008).

L'hypothèse neurodéveloppementale de la schizophrénie permet en outre d'intégrer un certain nombre d'observations et de considérer des facteurs de risque génétiques et non génétiques, ayant eu un impact sur le développement cérébral au cours de la période prénatale ou périnatale, ce qui semble plus réaliste que des modèles purement génétiques ou environnementaux. Parmi les facteurs non génétiques de nombreuses études ont montré une association entre certaines complications obstétricales et une augmentation du facteur de risque de schizophrénie (Brown & Derkits 2010; Byrne et al 2007; Cannon et al 2002), ces complications étant associées à des perturbations du neurodéveloppement (Marenco & Weinberger 2000; Rapoport et al 2005; Weinberger 1987). Parmi les mieux documentés figurent notamment le cas des infections virales *in utero* (Brown & Derkits 2010) ou les épisodes d'hypoxie à l'accouchement (Cannon et al 2002).

La contribution des facteurs génétiques dans l'hypothèse « neurodéveloppementale » serait liée à une perturbation du programme génétique normal de formation des synapses et de migration des neurones au cours du développement cérébral. De nombreux facteurs génétiques ont montré leur rôle dans les maladies neurodéveloppementales telles que l'autisme (Guilmatre et al 2009; Pinto et al 2010; Sebat et al 2007), et pour la schizophrénie, l'étude des gènes candidats identifiés et des anomalies structurales chromosomiques montre aussi une surreprésentation de gènes du neurodéveloppement en général (Guilmatre et al 2009; Sun et al 2010; Walsh et al 2008) et de la synapse glutamatergique en particulier (Banerjee et al 2010; Hahn et al 2006; Harrison & Weinberger 2005; Hayashi-Takagi & Sawa 2010; Pitcher et al 2011; Ross et al 2006). Le rôle crucial du système glutamatergique et des récepteurs NMDA dans un certain nombre d'étapes du neurodéveloppement est un argument supplémentaire en faveur de son implication dans la schizophrénie.

IV-Les hypothèses étiologiques : gènes ou environnement ?

A-Les études familiales

Il existe une augmentation de la prévalence de la schizophrénie chez les apparentés de sujets atteints comparée à celle de la population générale. Cette augmentation permet d'expliquer l'agrégation familiale de la maladie. Cependant, le mode de transmission ne correspond pas au profil de transmission mendélien monogénique. Ainsi le risque varie en fonction du lien de parenté et, il apparaît, lorsque

l'on considère l'ensemble des familles, que le risque de schizophrénie pour un apparenté d'un sujet atteint diminue plus rapidement que la proportion de matériel génétique partagé (**Figure 1**). Le risque de récurrence est lié aussi au nombre de sujets atteints dans la famille et à la sévérité du tableau clinique. Par ailleurs, ce sont plutôt les formes sporadiques de la maladie qui prédominent dans la population générale. En effet, il a été démontré qu'un déterminisme polygénique pour une maladie fréquente avec une prévalence à 1% et une forte héritabilité résulterait en une prédominance des formes sporadiques sur les formes familiales (Yang et al 2010). L'interprétation de ces études d'épidémiologie génétique est néanmoins influencée par la variation de l'expression phénotypique de la maladie.

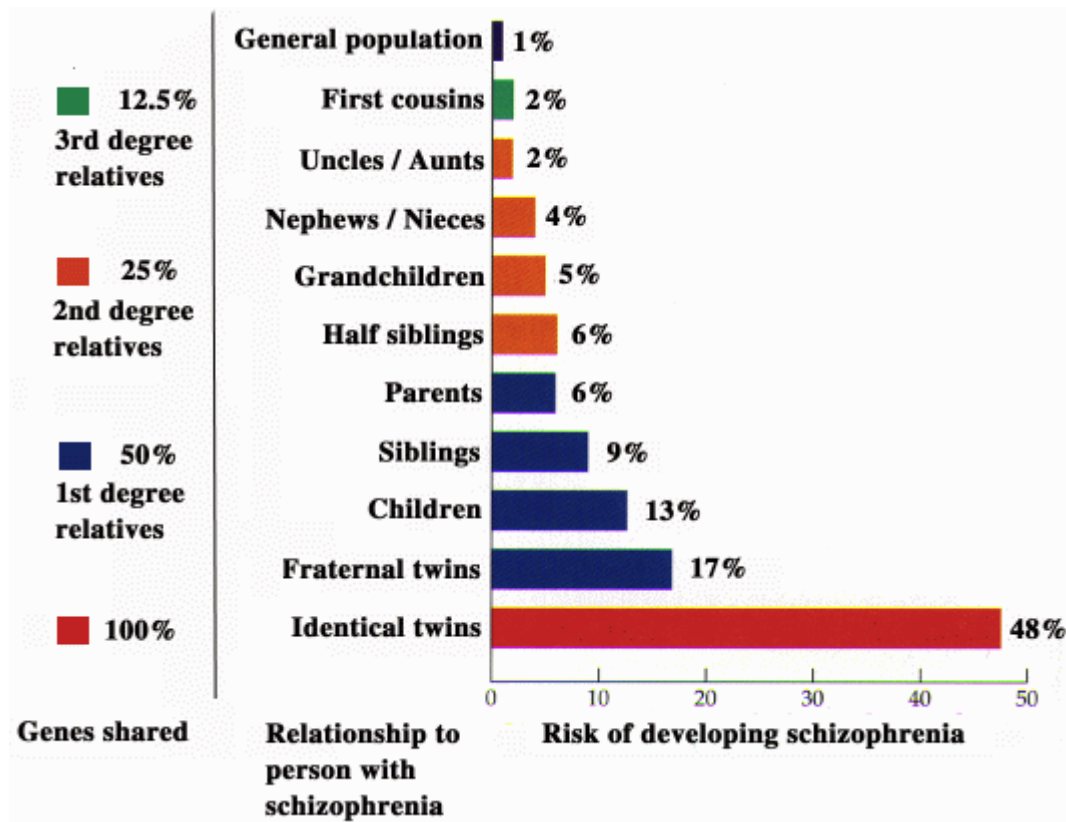


Figure 1 : Héritéité et risque de schizophrénie

(Adapté de Gottesman et al., Schizophrenia genesis : The origins of madness, 1991)

B-Les études de jumeaux

Les études de jumeaux monozygotes et dizygotes permettent de fournir une indication sur l'héritabilité d'une maladie. Ainsi, les différences entre jumeaux monozygotes sont attribuées à l'environnement et celles entre jumeaux dizygotes aux facteurs environnementaux et génétiques. L'héritabilité est une estimation de la proportion de la variance de la maladie expliquée par les facteurs génétiques. Elle a été estimée à environ 80% pour la schizophrénie (Sullivan et al 2003).

Le taux de concordance est une notion très utilisée et correspond à la probabilité du second jumeau de développer la maladie portée par le premier jumeau. Les taux de concordance pour la schizophrénie chez les jumeaux monozygotes sont estimés entre 40 et 50% contre 10 à 15% chez les jumeaux dizygotes (Sullivan et al 2003).

Pour des jumeaux monozygotes discordants pour la maladie, c'est-à-dire que l'un est atteint et l'autre non, le risque de troubles schizophréniques chez la descendance est similaire (Kringlen & Cramer 1989), ce qui suggère que le jumeau non atteint est porteur du risque génétique et héritable pour la schizophrénie sans l'exprimer forcément, reflétant l'implication très probable de facteurs environnementaux non partagés par les jumeaux. Récemment, des différences au niveau de la méthylation de l'ADN, marqueurs de la régulation épigénétique, ont été proposées afin d'expliquer la discordance des jumeaux monozygotes (Mill et al 2008).

C-Les études d'adoption

Ces études ont pour but d'évaluer séparément l'influence des facteurs génétiques des facteurs environnementaux. Ainsi, on étudie des enfants nés de parents atteints de schizophrénie et placés en adoption dans d'autres familles afin de déterminer si le risque de troubles schizophréniques reste identique à celui des parents biologiques ou devient celui des parents adoptifs. Ces études ont montré que ce risque restait élevé indépendamment de l'éclosion de la psychose chez les parents biologiques, avant ou après l'adoption, et indépendamment aussi du fait que les enfants adoptés ont été placés en famille ou en institution (Higgins 1966; Kety et al 1976). Il a été également montré, chez des enfants nés de mères atteintes, que le risque de troubles schizophréniques est le même qu'ils soient élevés par leur mères biologiques ou d'adoption (Higgins 1976). Aussi, il n'y avait pas d'augmentation de risque pour des enfants nés de mères non atteintes et élevés par des mères adoptives malades (Wender et al 1974). Ces données confirment que le risque de schizophrénie est transmis biologiquement plutôt que par le contexte familial adoptif et donc l'importance des facteurs génétiques quand au risque de développer la maladie.

D-Paradoxe évolutionniste

La schizophrénie est associée à une baisse de la fertilité, due probablement au handicap social. Cette hypofertilité étant plus marquée chez les hommes (Haukka et al 2003; Svensson et al 2007), ce phénomène tendrait à être accentuée par l'augmentation

de l'âge des hommes au moment du mariage dans les sociétés occidentales, l'âge d'apparition de la maladie restant stable.

La sélection naturelle devrait contribuer à diminuer la prévalence dans la population des variants géniques liés à la maladie. Cependant, la prévalence de la schizophrénie reste élevée et supérieure à celles des maladies mendéliennes. C'est le paradoxe évolutionniste. Plusieurs hypothèses ont été proposées (Keller & Miller 2006). Une augmentation de la fertilité chez les apparentés a été évoquée mais non confirmée par les études. Ces données infirment les hypothèses alternatives comme l'avantage des hétérozygotes ou le modèle de la pléiotropie antagoniste dans lequel un allèle qui diminue les chances de procréation pour un trait l'augmente pour un autre trait associé.

Une autre explication est que le phénotype clinique pourrait être très peu corrélé avec la vulnérabilité génétique sous-jacente, et que des phénotypes intermédiaires pourraient constituer un meilleur index reflétant les effets génétiques en comparaison à l'approche classique. Cependant, aucun des phénotypes intermédiaires candidats n'a montré à ce jour qu'il avait une héritabilité supérieure à celle du phénotype clinique (Allen et al 2009).

Enfin, les allèles augmentant le risque génétique de schizophrénie pourraient être maintenus dans la population générale contre toute sélection négative par le moyen d'un taux élevé de mutations. De nombreux résultats récents, y compris les travaux présentés dans cette thèse semblent actuellement étayer cette hypothèse (voir chapitre : **recherche de mutations rares**). L'âge paternel avancé pourrait être un facteur de risque qui agirait à travers l'augmentation du nombre de divisions des spermatogonies, faisant augmenter la chance d'apparition de nouvelles mutations.

Deuxième partie : Les facteurs génétiques dans la schizophrénie

I - Définitions

SNP = single nucleotide polymorphism = polymorphisme de nucléotide unique. Cela désigne la variation (polymorphisme) d'une seule paire de bases du génome, entre individus d'une même espèce. Ces variations sont très fréquentes (environ 1/1000 paire de bases dans le génome humain). Les SNP représentent 90 % de l'ensemble des variations génétiques humaines, et des SNP avec une fréquence allélique de ≥ 1 % sont présents tous les 100 à 300 paires de bases en moyenne dans le génome humain qui contiendrait 1.4 millions de SNP. Les SNP peuvent se retrouver au sein de régions codantes de gènes (exon), de régions non codantes de gènes (intron), ou de régions intergéniques. Dans le cas où les SNP se retrouvent au sein des régions codantes, ceux-ci ne vont pas obligatoirement modifier la séquence d'acide aminé de la protéine produite, et ce, grâce à la redondance du code génétique. Les SNP qui se retrouvent dans des régions non codantes peuvent avoir des conséquences sur l'épissage, les facteurs de transcription, ou sur les séquences d'ARN non codant.

Études de liaison : La recherche d'une liaison génétique a pour but de localiser une prédisposition génétique en identifiant des sites polymorphiques du génome dont les variations alléliques se transmettent avec la maladie dans les familles qui ont plus d'un membre atteint. Il s'agit d'une méthode de localisation d'un gène responsable de la maladie à un locus spécifique d'un chromosome donné (voir **figure 2**).

Les analyses statistiques consistent à déterminer si deux allèles d'un locus sont transmis ensemble à la descendance plus fréquemment que ne le voudrait le hasard. Cette approche peut être un outil puissant de détection des régions de susceptibilité dans les maladies monogéniques mais est mal adaptée pour les maladies génétiques complexes, le signal étant d'autant plus diffus que le nombre de gènes impliqués est important. L'identification d'un signal statistiquement significatif ou suggestif permet d'identifier une région génomique particulière qui peut être assez large et contenir plusieurs gènes.

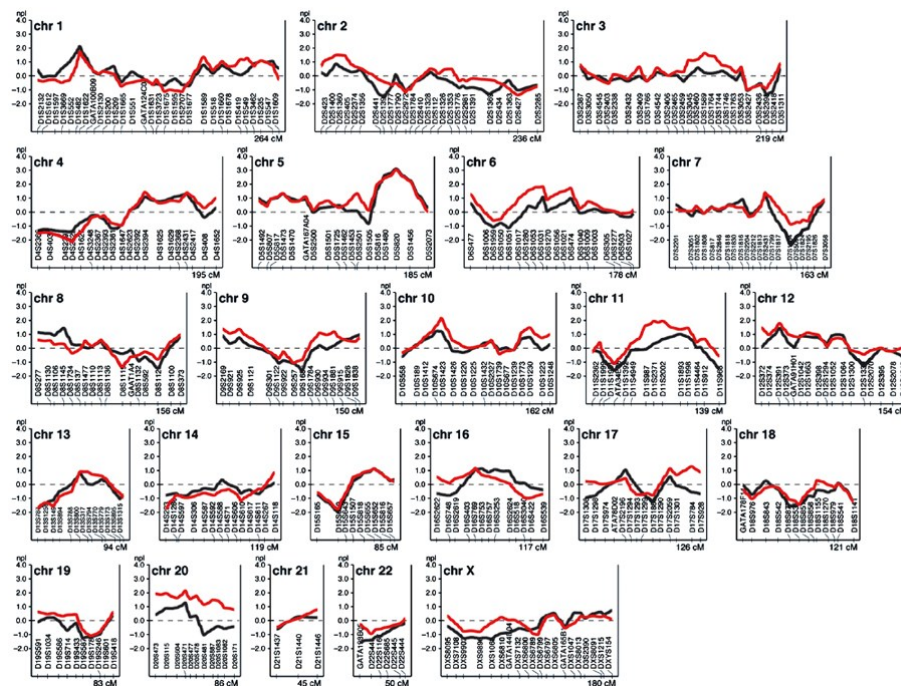


Figure 2 : Illustration des résultats d'une étude de liaison

Pour chaque chromosome, l'axe des abscisses indique les positions des marqueurs microsatellites. Les scores de liaison sont indiqués en rouge pour le phénotype « psychose » et en noir pour la schizophrénie. Une liaison est déclarée significative lorsque le score est supérieur ou égal à 3. (Tiré de Sklar et al. Mol Psychiatry 2004) (Sklar et al 2004).

Étude d'association génétique : Elles ont pour but de comparer la fréquence d'un (ou plusieurs) polymorphisme(s) entre des sujets cas et des sujets témoins dans une même population et identifier ainsi des gènes de prédisposition pouvant être associés à la maladie (voir **figure 3**). Depuis 2005, les puces permettent de séquencer un très grand nombre de SNP (jusqu'à 1 million), répartis sur l'ensemble du génome ou dans des régions sélectionnées, chez de nombreux individus. Ces puces sont devenues accessibles à des coûts raisonnables permettant l'émergence des **études d'association pangénomique** (GWAS : Genome wide association studies) en tant qu'outil pour décrypter les bases moléculaires de la schizophrénie. Ce type d'étude a déjà permis d'identifier des gènes candidats pour d'autres maladies complexes telles que le diabète de type 2 ou la sclérose en plaques (McCarthy et al 2008). Les études d'association qu'elles étudient des gènes candidats ou pangénomiques reposent sur l'hypothèse « variants fréquents-maladie fréquente ». Cette hypothèse suggère que la maladie résulterait de variants fréquents, chacun ayant un effet faible individuellement mais contribuant en interaction à l'apparition de la maladie. Une liste complète de l'ensemble des études d'association pangénomiques publiées est disponible sur le site du National Human Research Genome Institute (<http://www.genome.gov/26525384>).

Alternativement, l'association génétique peut être mesurée par le moyen d'études intra-familiales en analysant un polymorphisme chez un sujet atteint et ses deux parents. Chacun des 4 allèles parentaux de ce marqueur est défini soit comme transmis à l'enfant atteint, soit comme non-transmis. Le test (TDT) assume qu'un parent hétérozygote pour un allèle A1 associé à la maladie et un allèle A2 non associé à la maladie, transmettra plus souvent l'allèle A1 que l'allèle A2 à son enfant malade. La fréquence avec laquelle

les allèles A1 et A2 sont transmis ou non transmis aux enfants atteints sera donc différente de la fréquence attendue en cas de transmission mendélienne (50%, 50%).

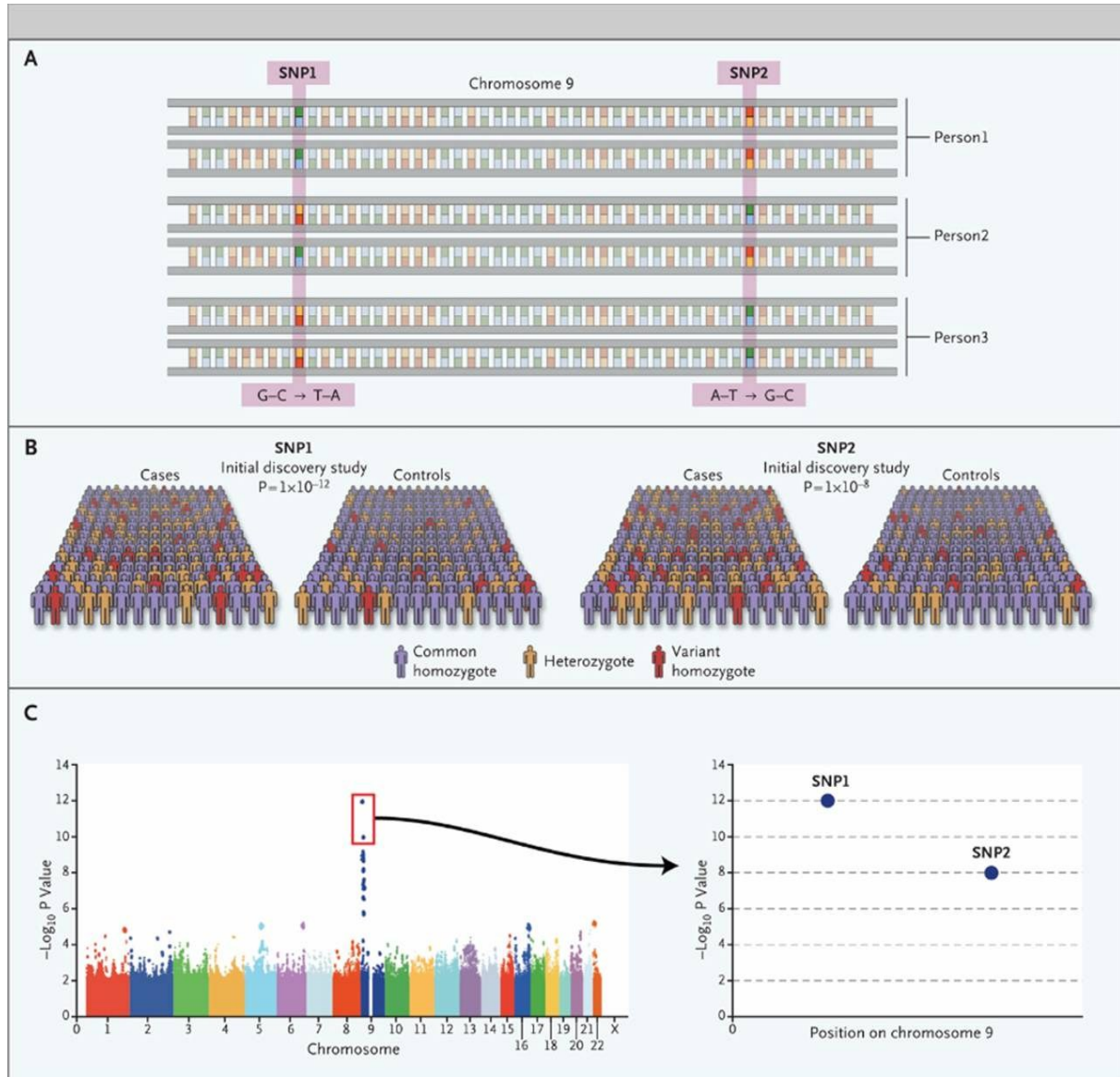


Figure 3 : Illustration du principe des études d'association pangénomique.

Les GWAS, tout comme les études d'association sont des études basées sur la comparaison entre un groupe de cas et un groupe de témoins. De très nombreux polymorphismes (Single nucleotide polymorphisms : SNP) (exemple en partie A) répartis sur l'ensemble du génome sont génotypés. La force de l'association est calculée par la comparaison de la prévalence allélique entre patients et contrôles (partie B). Dans cet exemple, SNP1 et SNP2, situés sur le chromosome 9, sont associés à la maladie avec des p-values de 10^{-12} et 10^{-8} (partie C).

(Tiré de Manolio et al., New England Journal of Medicine, 2010). (Manolio 2010)

Variants du nombre de copies : Copy number variations (CNV) : Ce sont des délétions ou des duplications de fragments d'ADN, allant de quelques centaines à plusieurs milliers de bases. De telles variations peuvent englober un ou plusieurs gènes, conduisant à un déséquilibre des niveaux d'expression de ces gènes, générant une diversité génétique qui n'a pu être explorée que grâce aux progrès des techniques de biologie moléculaire et l'apparition de puces permettant de scanner l'ensemble du génôme. Cela a conduit à une augmentation significative du nombre de CNV identifiés ces dernières années (Iafrate et al 2004; Sebat et al 2004).

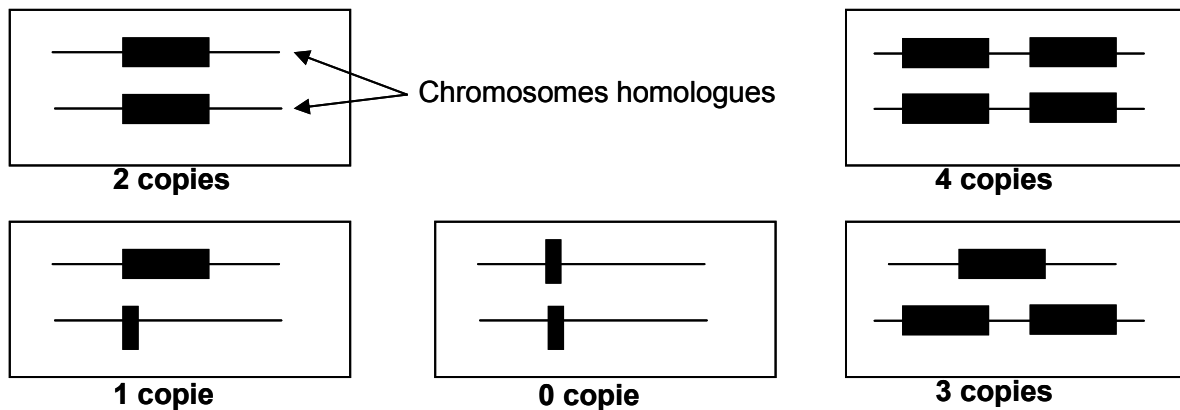


Figure 4 : Représentation de la diversité génétique générée par les CNV.

II- La recherche des marqueurs moléculaires

A-Études de liaison

Les premières études de liaison utilisaient un nombre limité de marqueurs polymorphes (quelques centaines de microsatellites) et produisaient des résultats discordants. Il existe actuellement plus d'une vingtaine d'études plus récentes dites « pangénomiques » car utilisant un nombre très élevé de marqueurs (des milliers de marqueurs de type SNP). Trois méta-analyses ont conclu que les régions 1q, 2q, 3p, 4q, 6p, 8p, 11q, 13q, 14p, 20q et 22q sont les régions les plus fréquemment retrouvées en liaison avec la schizophrénie (Badner & Gershon 2002; Lewis et al 2003; Ng et al 2009a). Cependant, seule la région 8p était confirmée par les trois méta-analyses. Il se pourrait en réalité qu'il n'existe pas un signal suffisamment puissant pour permettre d'incriminer une région de vulnérabilité génétique et qu'il n'existe pas de critères pour distinguer ce signal du bruit de fond (Crow 2007). De plus, dans l'éventualité de l'hypothèse de plusieurs gènes à petits effets et d'une grande hétérogénéité entre les populations, des effectifs très importants sont probablement nécessaires pour détecter un signal témoignant d'une liaison valable.

B-Études d'association, approche gènes candidats

Des centaines de gènes candidats ont été exploré faisant de la schizophrénie la maladie la plus étudiée par ce type d'approche. La principale critique formulée à l'encontre de ces études est qu'elles ne permettent de trouver que ce que l'on cherche : elles confirment, ou non, ce que l'on sait déjà ou ce que l'on avait des arguments raisonnables de suspecter. Si

les résultats ont généralement du sens, en permettant de consolider ou d'affiner des hypothèses, compte tenu des inconnues concernant la schizophrénie, il est vraisemblable qu'une telle démarche passe à côté de facteurs génétiques potentiellement importants. Le manque de puissance statistique, très fréquemment observé dans ces études, fait que la réplication des résultats initiaux est très rare, et, à ce jour, aucun gène candidat n'a été confirmé d'une manière claire et convaincante. L'inclusion d'effectifs de taille réduite, l'utilisation de techniques statistiques inadéquates ou de critères permissifs sont également des éléments générateurs de résultats non répliqués. Les analyses multiples et la sélection de données participent aussi à l'inflation des résultats faussement positifs. Il est également possible que l'hétérogénéité génétique empêcherait la réplication de certains résultats mais cet argument ne rend probablement pas compte de l'ensemble des données publiées. Enfin, l'existence d'épistasie et le non contrôle des facteurs environnementaux pourraient également expliquer cette non-réplication.

C-Études d'association pangénomique

Depuis 2006, au moins 12 études d'association pangénomique (voir **figure 5**) ont été entreprises afin d'identifier des gènes candidats pour la schizophrénie. Les premières incluant des effectifs relativement faibles (- de 2000 individus) (Kirov et al 2009b; Lencz et al 2007; Mah et al 2006; Shifman et al 2008; Sullivan et al 2008), et incluant, tout du moins pour la première d'entre elles (Mah et al 2006), un nombre restreint de SNP (~25000) analysés ce qui a limité la puissance de l'analyse. Les plus récentes incluant un très grand nombre d'individus (entre 10 000 et 20 000) et jusqu'à 1 million de SNP

(Athanasiau et al 2010; Ikeda et al 2011; Need et al 2009; O'Donovan et al 2008; Purcell et al 2009; Shi et al 2009; Stefansson et al 2009).

Au final, les études d'association pangénomique nous indiquent que même s'il y a de nombreux et faibles signaux répartis sur l'ensemble du génome, seules quelques régions chromosomiques sont associées de façon significative avec la schizophrénie et répliquées de façon consistante. Cependant, ces régions chromosomiques n'ont qu'un effet mineur avec un risque relatif (Odds ratio : OR) inférieur à 1.2, d'autre part l'effet fonctionnel de la vaste majorité des SNP identifiés n'a pas été exploré ni identifié. De plus, même considérées dans leur ensemble, ces mutations ne permettent de caractériser qu'une faible partie de l'héritabilité de la maladie (probablement moins de 2%) (Crow 2011; Purcell et al 2009). Enfin, ce type d'analyse suppose que les différences des fréquences alléliques entre les cas et les témoins sont uniquement liées au fait d'être porteur de la maladie et non à des différences liées à des origines ethniques, associées à des fréquences différentes. Cette différence génétique entre les sujets cas et les sujets témoins, appelée stratification de la population, est un biais méthodologique sérieux pour cette approche.

Il n'en reste pas moins que nombre de ces gènes sont probablement impliqués dans l'étiologie de la maladie, conférant une susceptibilité à la maladie relativement faible, et, d'une manière générale, les études d'association ont permis à la fois l'identification de nouveaux gènes de susceptibilité et ouvert la voie à une compréhension multifactorielle du risque génétique associé à la schizophrénie avec probablement plus de 1000 gènes impliqués.

Étude	Taille de l'échantillon		Nombre de SNP testés	Gène avec SNP significatif
	Cas	Contrôles		
Mah et al, 2006(Mah et al 2006)	520	555	25 494	<i>PLXNA2</i>
Shifman et al, 2008(Shifman et al 2008)	1 009	3 158	510 552	<i>RELN</i>
Kirov et al, 2009(Kirov et al 2009b)	574 trios	605 trios	433 680	<i>CCDC60, RBP1</i>
Lencz et al, 2007(Lencz et al 2007)	249	175	439 511	<i>CSF2RA, IL3RA</i>
Sullivan et al, 2008(Sullivan et al 2008)	738	733	492 900	/
O'Donovan et al, 2008(O'Donovan et al 2008)	7 145	12 834	362 532	<i>ZNF804A</i>
Need et al, 2009(Need et al 2009)	1 734	14 455	312 565	<i>ADAMTSL3</i>
Purcell et al, 2009(Purcell et al 2009)	11 330	22 664	739 995	<i>MYO18B, NOTCH4, RPP21</i>
Stefansson et al, 2009(Stefansson et al 2009)	7 662	29 053	314 868	<i>MHC^I region, NRG1, PRSS16, TCF4, PGDB1, NOTCH4</i>
Shi et al, 2009(Shi et al 2009)	11 975	22 703	696 788	<i>MHC^I region</i>
Athanasios et al, 2010(Athanasios et al 2010)	2 867	14 085	572 888	<i>PLAA, ACSM1, ANK3</i>

et al 2010)				
Ikeda et al, 2010(Ikeda et al 2011)	2 565	5 953	297 645	<i>SULT6B1, NOTCH4</i>

Figure 5 : Résumé des études d'association pangénomique de la schizophrénie
<http://www.genome.gov/26525384>

¹ : Major Histocompatibility complex

D-Étude des anomalies chromosomiques et les variants du nombre de copie

Les variants du nombre de copies (CNV : copy number variants) sont des délétions ou des duplications de fragments génomiques de taille variables pouvant aller de 1kb à plusieurs mégabases. L'étude de ces anomalies chromosomiques chez les patients atteints de schizophrénie a été une des premières stratégies employées pour décrypter la contribution génétique de la maladie (Blackwood et al 2008; Shastri 1999). Les premières régions identifiées ont été, du fait des limitations technologiques, de grande taille ce qui a rendu difficile l'identification des gènes candidats. Cependant, la résolution de ces études ayant constamment été améliorée, allant de 2.3Mb en 2006 à moins de 100kb en 2010 en partie grâce à la ré-analyse des données des études d'association pangénomiques, cela a permis d'identifier des délétions de taille inférieure, et ce d'autant plus qu'il semble qu'il y ait une relation inverse entre la fréquence et la taille des CNV, et une majorité des CNV auraient une taille inférieure à 50kb (Korbel et al 2007).

La délétion en 22q11, aussi connue comme la région du syndrome vélo-cardio-facial/DiGeorge est caractérisée par une perte de 1.5 à 3 mégabases sur l'un des deux chromosomes. Elle fait partie des régions identifiées le plus précocement (du fait de sa taille) et avec les meilleurs arguments en faveur d'une implication dans la schizophrénie. Outre la présence chez 30% des porteurs de la délétion de symptômes psychotiques, souvent indissociables des symptômes de la schizophrénie, la délétion est associée à un risque 25 fois supérieur de développer la schizophrénie ainsi que de développer d'autres maladies neuropsychiatriques (Lindsay et al 1995; Murphy et al 1999). Parmi les gènes contenus dans la région, les gènes de la catechol-O-méthyltransférase (*COMT*) et de la proline déshydrogénase (*PRODH*) ont été les plus étudiés mais la contribution génétique aux symptômes associés à cette délétion reste à être mieux définie (Jacquet et al 2002; Karayiorgou et al 1995; Williams et al 2007).

Parmi les anomalies chromosomiques les plus précocement associées à la schizophrénie, on peut citer aussi la translocation équilibrée t(1;11)(q42.1;q14.3), identifiée pour la première fois dans une grande famille écossaise et co-ségréant avec de nombreux individus atteints que ce soit de schizophrénie, de troubles bipolaires ou de dépression majeure (St Clair et al 1990). Cette translocation interrompt le gène *Disrupted in schizophrenia DISC1* ainsi que le gène *DISC2* (Millar et al 2000) mais les mécanismes physiopathologiques restent à être mieux établis. (Porteous et al 2006) Elle est associée à un risque 25 à 50 fois supérieur de développer l'un de ces troubles (Chubb et al 2008).

Les études plus récentes n'ont fait que confirmer l'implication des microdélétions et microduplications dans l'étiologie d'un certain nombre de maladies

neurodéveloppementales, permettant l'identification de régions chromosomiques toujours plus précises.

D'une manière générale, la fréquence des CNV est augmentée chez les patients (2008; Bassett et al 2010; Guilmatre et al 2009; Kirov et al 2009a; Ng et al 2009a; O'Donovan et al 2009; Purcell et al 2009; Sebat 2007; Stefansson et al 2008; Walsh et al 2008; Xu et al 2008; Xu et al 2009) mais l'amplitude de l'augmentation varie grandement selon la taille et la fréquence des CNV étudiés, les CNV rares (présents dans moins de 1% des individus) étant plus significativement associés à la schizophrénie (2008; Kirov et al 2009a; Walsh et al 2008; Xu et al 2009). Ces CNV ont des risques relatifs associés très supérieurs à ceux des variants identifiés par les études d'association. L'observation de ces variations a conduit à une reconsidération de l'hypothèse des « variants fréquents ».

L'étude de Walsh (Walsh et al 2008) en particulier a montré l'implication de 53 CNV rares présents plus fréquemment chez les patients (15%), notamment chez ceux ayant eu un âge de début précoce (20%). Bien que la plupart des mutations soient présentes chez les contrôles, suggérant que ces mutations ne sont pas très pénétrantes, les résultats suggèrent aussi que certaines localisations seraient plus à risque. Alors que les CNV des témoins semblent répartis aléatoirement, les CNV des patients toucheraient de façon préférentielle des réseaux de gènes impliqués dans la synapse et le neurodéveloppement, cela inclut celui du récepteur de la Neuregulin (ERBB4), des gènes de la voie ERK/AMPK, des gènes impliqués dans la potentialisation à long terme, le guidage axonal, la transmission glutamatergique (Guilmatre et al 2009; Kirov et al 2008; Mei & Xiong 2008; Rujescu et al 2009; Walsh et al 2008). Bien que ces études ne s'accompagnent généralement pas d'une démonstration des effets fonctionnels de la

mutation, il est raisonnable de penser que certaines de ces anomalies puissent avoir un effet important sur le risque de schizophrénie, voire un risque causal.

Néanmoins, tout comme pour les variants identifiés par les études d'association pangénomiques, les CNV les plus fréquemment observés, même considérés collectivement, ne permettraient d'expliquer qu'une faible proportion des cas. De plus, pour la plupart de ces cas, le phénotype associé n'est pas spécifique à la maladie mais augmente parallèlement le risque pour d'autres pathologies, en particulier l'autisme, les troubles bipolaires ou la déficience mentale (voir chapitre **problèmes nosographiques : vers un continuum SCZ BD ASD**).

La signification de cette nouvelle forme de variation génétique n'est pas totalement bien déterminée. Les CNV sont répartis sur l'ensemble du génome mais certaines localisations sont préférentielles, les régions riches en GC par exemple fournissent un terrain favorable aux événements de duplication segmentale pouvant conduire à ces variations. Les CNV peuvent jouer un rôle pathologique de différentes manières, s'ils ont pour conséquence la perte de fonction d'un gène causal. Alors ils aboutissent à une mutation récessive ou dominante selon le rôle du gène touché. Ils peuvent aussi augmenter ou diminuer le nombre d'allèle fonctionnel d'un gène conduisant à une variation de son niveau d'expression. Avec les mêmes conséquences, ils peuvent toucher un élément régulateur.

Enfin ils peuvent moduler l'expression phénotypique d'une autre mutation, comme cela a été observé dans le cas de l'amyotrophie spinale infantile. Cette maladie est due à des mutations homozygotes du gène SMN1 et présente plusieurs formes : une forme sévère, dite classique, et des formes beaucoup moins graves. Il a été démontré que ces différences s'expliquent par le nombre de copies du gène SMN2. Celui-ci a la même

fonction que SMN1 à la différence que son ARN transcrit est dégradé et, chez les patients qui ont suffisamment d'exemplaires du gène, la quantité de protéine SMN2 traduite est suffisante pour remplacer plus ou moins partiellement SMN1 d'où une expression phénotypique moins sévère de la maladie. Au final il apparaît que les CNV sont impliqués de façon cruciale dans la diversité génétique, contribuant plus que les SNP à la variation globale de séquence génomique (Conrad et al 2010; McCarroll et al 2008; Redon et al 2006; Shianna & Willard 2006).

Bien que la plupart des CNV soient transmis, une partie sont *de novo* (c'est-à-dire que ces CNV sont absents chez les parents) et sont associés de façon significative à la maladie, notamment chez un certain nombre de patients n'ayant pas d'histoire familiale de troubles psychiatriques (Ng et al 2009a; Xu et al 2008). La microdélétion en 22q11 par exemple est observée dans 90% des cas *de novo*

L'implication de mutations apparaissant *de novo* a encore été peu explorée mais pourrait fournir une explication plausible à la part manquante de l'hérabilité et fait l'objet de nos travaux. De façon intéressante, il a été montré dans une étude récente (Johansson & Feuk 2011) que les régions qui présente le moins de variabilité génétique en termes de CNV dans la population générale sont aussi celles qui contiennent les CNV *de novo* identifiées dans l'autisme ou le retard mental, renforçant l'idée que ces événements sont bien impliqués dans l'étiologie de ces maladies.

E-Recherche des mutations rares

La découverte de l'implication de certaines mutations rares et *de novo* grâce aux études des anomalies chromosomiques et des CNV a permis une nouvelle approche du risque génétique de la maladie, prenant en considération l'implication de telles mutations ayant une forte pénétrance (2008; Walsh et al 2008; Xu et al 2008). En plus de fournir de nouvelles pistes physiopathologiques de la maladie, l'implication de telles mutations permet d'expliquer une part de l'« héritabilité manquante » de la schizophrénie et un certain nombre d'observations cliniques.

La possibilité de réanalyser les données des études d'association pangénomique et celle d'obtenir de très grande quantité de séquences grâce aux progrès biotechnologiques (séquenceurs de 2^{ème} et de 3^{ème} génération) associées à la baisse des coûts du séquençage a ouvert la voie aux projets de séquençage à grande échelle ainsi qu'à l'identification de mutations ponctuelles rares et *de novo* associées à la maladie. L'accessibilité des nouvelles méthodologies de re-séquençage rend possible la recherche de mutation ponctuelle qui puisse avoir un effet majeur. C'est la démarche adoptée dans notre travail. Tout comme pour les CNV, un certain nombre d'études, dont celles du projet présenté ici ont déjà montré l'implication de tels variants dans l'autisme ou le retard mental.

F-Problèmes nosographiques : vers un continuum schizophrénie, troubles bipolaires, troubles du spectre autistique

La schizophrénie a été considérée comme une maladie distincte depuis plus d'un siècle mais sa proximité avec d'autres troubles, notamment les troubles bipolaires, a

d'emblée été évoquée. Les études génétiques, cognitives et épidémiologiques ont montré que les bases physiopathologiques de la schizophrénie dépassent probablement les frontières nosologiques classiques. Les études familiales ont montré une agrégation familiale du risque de schizophrénie et du risque de troubles bipolaires (Van Snellenberg & de Candia 2009). En plus de partager des caractéristiques épidémiologiques (prévalence, héritabilité, âges de début) comparables et un certain nombre de symptômes communs, la plupart des SNP et des anomalies chromosomiques identifiés par les études confèrent un risque supérieur à la schizophrénie, aux troubles bipolaires mais aussi à d'autres pathologies psychiatriques et une majorité des loci associés à la schizophrénie ont aussi été associés dans les troubles bipolaires, les troubles du spectre autistique et réciproquement (2008; Guilmatre et al 2009; Mefford et al 2008; Stefansson et al 2008; Williams et al 2011). A titre d'exemple, la délétion en 22q11, démontrée comme étant surreprésentée chez les patients atteints de schizophrénie (Lindsay et al 1995) l'est aussi chez les patients atteints de troubles de l'humeur (Green et al 2009), d'anxiété et de troubles de l'attention (Jolin et al 2011), de troubles envahissants du développement (Niklasson et al 2009) et à des niveaux variés de déficit intellectuels (Green et al 2009).

Ces observations posent le problème de la spécificité des mutations identifiées pour la schizophrénie et des biais méthodologiques liés à la sélection des patients. Par exemple la déficience intellectuelle est une pathologie associée de façon claire avec un certain nombre de variants rares, les CNV notamment. Or, le quotient intellectuel des patients atteints de schizophrénie est diminué de façon légère mais significative. Par conséquent, l'inclusion de contrôles n'ayant pas de déficience intellectuelle dans les

analyses pourrait induire l'identification de variants plus associés à cette condition qu'à la schizophrénie.

En réalité, les résultats s'accumulent ces derniers temps pour montrer l'existence d'un probable continuum avec une importante décroissance des anomalies du développement cérébral dans l'expression clinique entre retard mental et schizophrénie, en passant par l'autisme. Certaines études ont tenté de mettre en commun des individus atteints de ces différentes pathologies, afin d'augmenter la puissance de la recherche des marqueurs moléculaires, mais cela a globalement été peu informatif, et il devient important d'identifier des mutations qui soient spécifiques aux mécanismes biologiques de ces maladies. Cela nécessite une stratégie de réplique des résultats dans des cohortes de patients de pathologies distinctes et des diagnostics cliniques rigoureux tels que présentés dans ces travaux.

G- Approche endophénotypique :

Les endophénotypes ou phénotypes intermédiaires sont des caractéristiques associées à la maladie, qui pourraient refléter une influence génétique même en l'absence de toute manifestation clinique pathologique, et en particulier psychiatrique. Ils seraient donc détectables non seulement chez les sujets atteints mais aussi chez les sujets porteurs d'un risque génétique de développer la maladie comme les parents des sujets atteints. Il a été proposé que les endophénotypes ont une architecture génétique plus simple que le phénotype clinique lui-même ce qui simplifierait l'étude génétique des maladies complexes tout en augmentant leur puissance (Gottesman & Gould 2003).

Plusieurs critères ont été proposés pour qu'une anomalie biologique ou comportementale soit considérée comme un endophénotype candidat. Un endophénotype doit être associé à la maladie, avoir une héritabilité importante et être présent à un degré plus important chez les sujets à risque génétique pour la maladie, comme les parents sains, en comparaison à la population générale. Il doit être également relativement indépendant des variations des symptômes cliniques et des facteurs d'environnement.

Dans la schizophrénie, des mesures électrophysiologiques, des phénotypes d'imagerie structurale et/ou fonctionnelle et des paramètres cognitifs et oculomoteurs ont été étudiés comme des endophénotypes candidats (Allen et al 2009). A ce jour aucun de ces endophénotypes ne remplit pleinement ces critères. Cependant, leur utilisation a permis d'identifier des gènes candidats et de mieux comprendre les voies fonctionnelles sous-jacentes. Par exemple, l'association du gène ZNF804A (codant pour la protéine en doigt de zinc 804A, de fonction inconnue) avec la schizophrénie a été confirmée par plusieurs études pangénomique GWAS. L'utilisation de l'approche endophénotypique a permis de montrer que ce gène était associé au fonctionnement cérébral durant une tâche de mémoire de travail (Esslinger et al 2009) qui est une fonction cognitive très importante et connue pour être déficitaire chez les patients atteints de schizophrénie.

Au total, l'existence de facteurs génétiques dans la schizophrénie est avérée mais leur nature précise et les modalités de transmission restent incertaines, et peut-être variables selon les formes de la maladie ou selon les familles. Toutefois, les arguments convergent pour suggérer que la schizophrénie, comme d'autres maladies neurodéveloppementales (autisme) et / ou cognitive (déficience mentale), est la conséquence d'une perturbation

des gènes impliqués au niveau synaptique, et en particulier dans la plasticité et la transmission synaptique. Le système glutamatergique pourrait être l'un des principaux mécanismes altérés.

Troisième partie : Glutamate et schizophrénie

I-Systèmes glutamatergiques

A-Introduction

Le glutamate est le neurotransmetteur le plus abondant du système nerveux central et le principal neurotransmetteur excitateur, utilisé par plus de 80 % des neurones du cerveau. Il joue un rôle crucial dans le développement cérébral, la survie et la différenciation cellulaire ainsi que dans la plasticité synaptique. La plupart des fonctions cérébrales normales telles que la cognition, la mémoire et l'apprentissage sont aussi dépendantes de ce neurotransmetteur.

Ironiquement, du fait qu'il soit ubiquitaire, le glutamate n'a été identifié que très tardivement en tant que neurotransmetteur et, bien qu'il soit présent à des concentrations élevées dans tout le système nerveux, 99.99% du glutamate est intracellulaire et les concentrations en glutamate extracellulaire doivent être extrêmement modulées en raison de ses propriétés excitotoxiques lorsqu'il est en excès.

Le glutamate est impliqué dans une large gamme d'effets complexes et variés au niveau neuronal et la spécialisation de ses effets est en partie expliquée par la grande variété de ses récepteurs, ainsi que par la spécificité de leur localisation subcellulaire.

B-Voies de synthèse du glutamate

Le glutamate est synthétisé dans les mitochondries des neurones pré-synaptiques à partir d'un précurseur, la glutamine, qui subit une désamination oxydative par l'action de l'enzyme glutaminase. Il existe une autre voie de synthèse utilisant l'alpha-cétoglutarate, un métabolite du cycle de Krebs mitochondrial, comme précurseur et la transaminase comme enzyme. Le glutamate est ensuite internalisé dans les vésicules synaptiques grâce à des transporteurs vésiculaires pour être libéré dans la synapse à l'arrivée d'un potentiel d'action. L'excès de neurotransmetteur dans la fente synaptique est ensuite recapté, quasi exclusivement par les astrocytes grâce à leurs transporteurs actifs d'acides aminés excitateurs. Le glutamate est alors aminé sous l'action de l'enzyme glutamine synthétase. La glutamine ainsi formée est relarguée par les astrocytes et finalement récupérée par les neurones glutamatergiques. C'est ce que l'on appelle le cycle glutamate-glutamine.

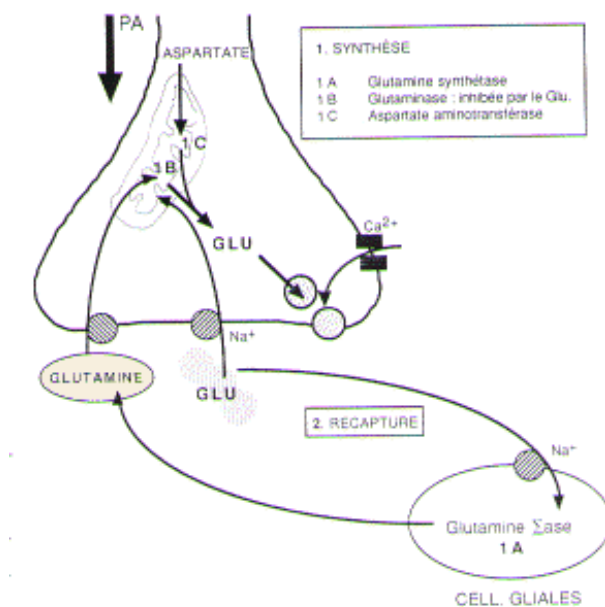


Figure 6 : Voies de synthèse du glutamate

C-Les récepteurs glutamatergiques

Plusieurs types de récepteurs ionotropiques et métabotropiques au glutamate ont été identifiés. Concernant les récepteurs ionotropiques, aussi appelés récepteurs canaux, ils sont regroupés en 3 classes : NMDA (N-Méthyl-D-Aspartate), AMPA (α -amino-3-hydroxy-5-methyl-isoxazole-4-propionyl acid) et kainate. Si l'on se base sur des critères d'homologie de séquence, il semble qu'ils proviennent d'au moins 6 familles de gènes ayant une origine commune : une famille pour les récepteurs AMPA (*GRIA1-4*), deux pour les récepteurs kainate (*GRIK1-3*, *GRIK4-5*) et trois pour les récepteurs NMDA (*GRIN1*, *GRIN2A-D*, *GRIN3A-B*) (voir **figure 7**). Ces gènes codent pour des sous-unités qui vont former un récepteur fonctionnel sous une forme hétérotétramérique.

Les récepteurs métabotropiques sont des récepteurs couplés à des protéines G qui modulent la transmission glutamatergique. Les récepteurs mGluR3 par exemple inhibent la recapture du glutamate (Xi et al 2002).

Receptor Family	Subunit	Gene	Chromosome (human)
AMPA	GluR ₁	<i>GRIA1</i>	5q33
	GluR ₂	<i>GRIA2</i>	4q32-33
	GluR ₃	<i>GRIA3</i>	Xq25-26
	GluR ₄	<i>GRIA4</i>	11q22-23
Kainate	GluR ₅	<i>GRIK1</i>	21q21.1-22.1
	GluR ₆	<i>GRIK2</i>	6q16.3-q21
	GluR ₇	<i>GRIK3</i>	1p34-p33
	KA-1	<i>GRIK4</i>	11q22.3
	KA-2	<i>GRIK5</i>	19q13.2
NMDA	NR1	<i>GRIN1</i>	9q34.3
	NR2A	<i>GRIN2A</i>	16p13.2
	NR2B	<i>GRIN2B</i>	12p12
	NR2C	<i>GRIN2C</i>	17q24-q25
	NR2D	<i>GRIN2D</i>	19q13.1qter
	NR3A	<i>GRIN3A</i>	9q31.1
	NR3B	<i>GRIN3B</i>	19p13.3

Figure 7 : Récepteurs ionotropiques au glutamate
(Adapté de Waxman, Molecular neurology, 2007)

D- Synapse glutamatergique et plasticité synaptique.

La synapse glutamatergique réunit le bouton présynaptique, la dendrite postsynaptique et le pied astrocytaire.

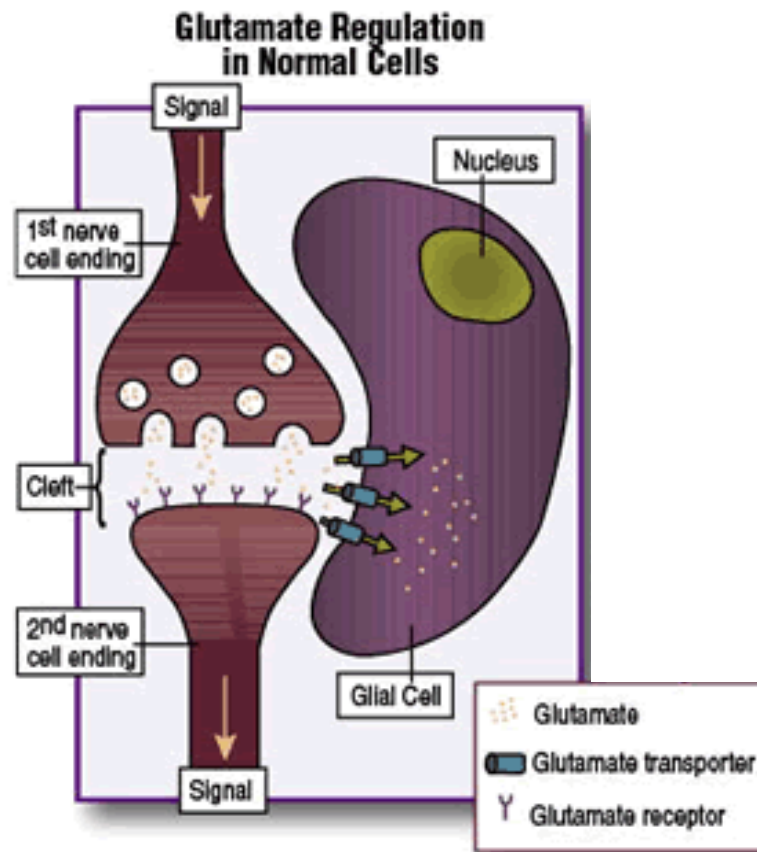


Figure 8 : La synapse glutamatergique

L'arrivée de potentiels d'action au niveau de la terminaison présynaptique entraîne la libération du glutamate dans la fente synaptique. Celui-ci se lie à ses récepteurs présents au niveau de la membrane postsynaptique. L'activation des récepteurs AMPA et kainate permet l'entrée de sodium, initiant la dépolarisation postsynaptique. Le changement du potentiel de membrane permet le relargage de l'ion magnésium bloquant les récepteurs NMDA. L'entrée de calcium au travers des récepteurs NMDA déclenche des cascades de

signalisation conduisant à la synthèse de protéines dont certaines ont pour effet l'augmentation prolongée de l'efficacité de la synapse par des changements morphologiques neuronaux durables (Bading et al 1993). Les récepteurs kainate présynaptiques pour leur part semblent faciliter la transmission synaptique en augmentant la libération de neurotransmetteurs.

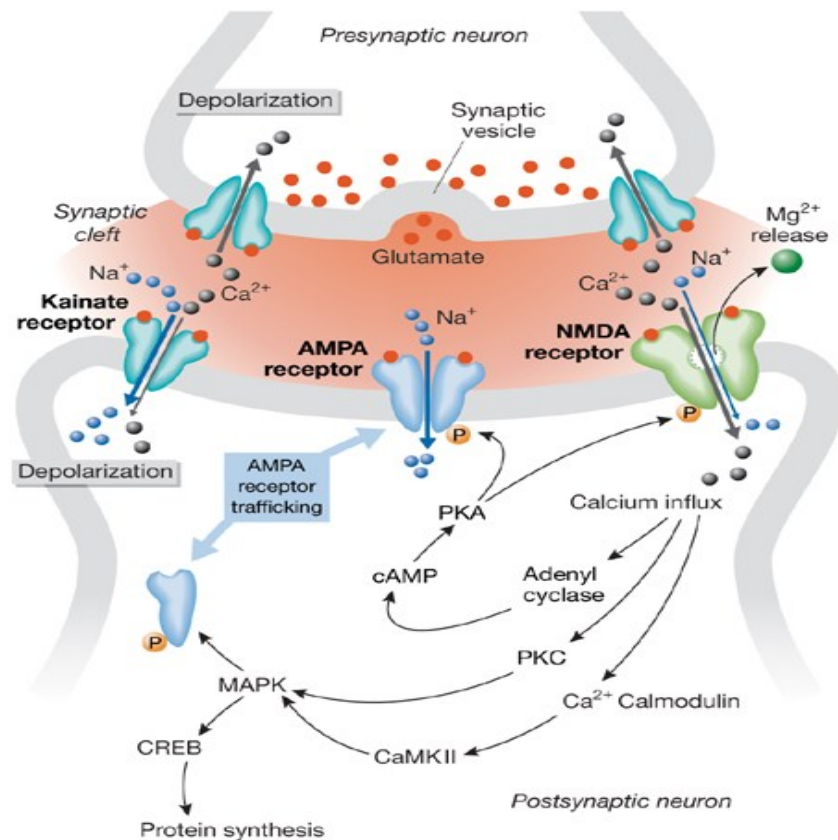


Figure 9 : Récepteurs glutamatergiques et plasticité synaptique

AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazolepropionate); CaMKII, (calcium/calmodulin-dependent kinase II); CREB (cAMP response element binding protein); MAPK (mitogen-activated protein kinase); PKA (protein kinase A); PKC (protein kinase C)

(Tiré de Voglis et al., 2006)(Voglis & Tavernarakis 2006)

Les synapses évoluent avec le temps, certaines disparaissant, d'autres se créant mais toutes se modifiant et tendant, soit à renforcer, soit à affaiblir, la communication entre deux neurones.

La plasticité synaptique désigne l'ensemble des modifications morphologiques, chimiques et fonctionnelles qui interviennent au cours du temps au niveau de la synapse.

La plasticité synaptique est à la base des processus d'apprentissage et de mémorisation à long terme (Collingridge et al 2004). Ces mécanismes sont sous-tendus au niveau cellulaire par la *potentialisation à long terme* (LTP) et la dépression à long terme (LTD). Ceux-ci désignent le renforcement ou l'affaiblissement durable des synapses entre deux neurones qui sont activés simultanément et ont été découverts en 1973, initialement dans l'hippocampe mais il a depuis été montré que ces mécanismes ont lieu dans de nombreuses régions du cortex.

Les récepteurs glutamatergiques et les récepteurs NMDA en particulier sont des acteurs clefs dans la plasticité synaptique à court terme et peuvent à ce titre exercer un effet d'inducteur de changements d'efficacité synaptique à plus long terme.

Les récepteurs NMDA sont également les principaux médiateurs du phénomène d'*excitotoxicité* (Waxman and Lynch 2005). L'activation excessive de ces récepteurs lors d'un relargage massif de glutamate entraîne une entrée massive de calcium dans la cellule et ainsi la mort cellulaire. Ce mécanisme physiopathologique survient dans plusieurs pathologies neurologiques comme l'épilepsie, les accidents vasculaires cérébraux ou certaines maladies neurodégénératives (chorée de Huntington, maladie d'Alzheimer, maladie de Parkinson...) et pourrait aussi être impliqué dans l'autisme et la schizophrénie.

II- Récepteurs NMDA : Description et propriétés

Les récepteurs NMDAR sont des récepteurs voltage-dépendant. Une dépolarisation de la membrane plasmique aux alentours de -40 mV est nécessaire pour lever le blocage du canal par l'ion magnésium (Nowak et al 1984). L'activation du récepteur est dépendante de la fixation des deux coagonistes, la glycine et le glutamate. Les récepteurs NMDA sont considérés comme des détecteurs de coïncidence dans la mesure où ils permettent l'entrée de calcium seulement lorsque les cellules pré et postsynaptiques sont activés. En plus de leur rôle dans la transmission synaptique excitatrice, cette propriété place les récepteurs du glutamate de type NMDA au centre de la potentialisation à long terme, de la dépression à long terme et de la plasticité synaptique en général (Liu et al 2004). Les récepteurs NMDA peuvent être pré ou postsynaptiques, participant à l'intégration du signal exciteur. Ils sont situés principalement au niveau des épines dendritiques mais ils sont présents aussi au niveau astrocytaire.

De par leur rôle dans la plasticité synaptique, les récepteurs NMDA sont impliqués dans les processus de mémoire, d'apprentissage, les fonctions cognitives. Au niveau synaptique et cellulaire, leur rôle a été bien établi grâce aux modèles animaux où ces récepteurs ont été modifiés génétiquement ou pharmacologiquement. Chez l'homme, ces rôles ont été mis en évidence de manière plus indirecte par l'utilisation de molécules antagonistes des récepteurs NMDA et l'étude de tissu cérébral. Il apparaît que la grande diversité des propriétés physiologiques et pharmacologiques des récepteurs NMDA pourrait reposer principalement sur la spécificité des sous-unités composant le récepteur

(Cull-Candy & Leszkiewicz 2004). Celle-ci varie selon la région cérébrale, l'activité et la période du développement (Ewald & Cline 2009).

Sept gènes codent pour des sous-unités du récepteur NMDA : *GRIN1*, *GRIN2A*, *GRIN2B*, *GRIN2C*, *GRIN2D*, *GRIN3A* et *GRIN3B*. Ces sous unités ont pour nom NR1, NR2A-D, NR3A et NR3B. Les récepteurs NMDA fonctionnels sont des tétramères composés de deux sous-unités NR1 en association avec deux autres sous-unités NR2 (A, B, C ou D) et/ou NR3 (A ou B). Toutes les sous-unités présentent une organisation fonctionnelle similaire avec un domaine N-terminal extracellulaire, un premier domaine transmembranaire (TM1), une boucle réentrante formant le deuxième domaine transmembranaire (TM2), un troisième domaine transmembranaire (TM3) suivi d'une boucle intracellulaire. Enfin le quatrième domaine transmembranaire (TM4) et le domaine C terminal intracellulaire. (Voir **figure 10**). Les domaines extracellulaires sont impliqués dans l'ancrage des ligands (Monaghan & Jane 2009) et possèdent de nombreux sites de fixation de modulateurs pharmacologiques (Williams 2009), les domaines transmembranaires et TM2 en particulier forment le canal sous forme tétramérique alors que les domaines intracellulaires C-terminaux sont impliqués dans l'ancrage aux protéines de la densité postsynaptique. Ils sont aussi la cible de modifications post-translationnelles fréquentes, incluant phosphorylations, glycosylations, ubiquitinations pouvant modifier leurs localisations et leurs fonctions (Salter et al 2009).

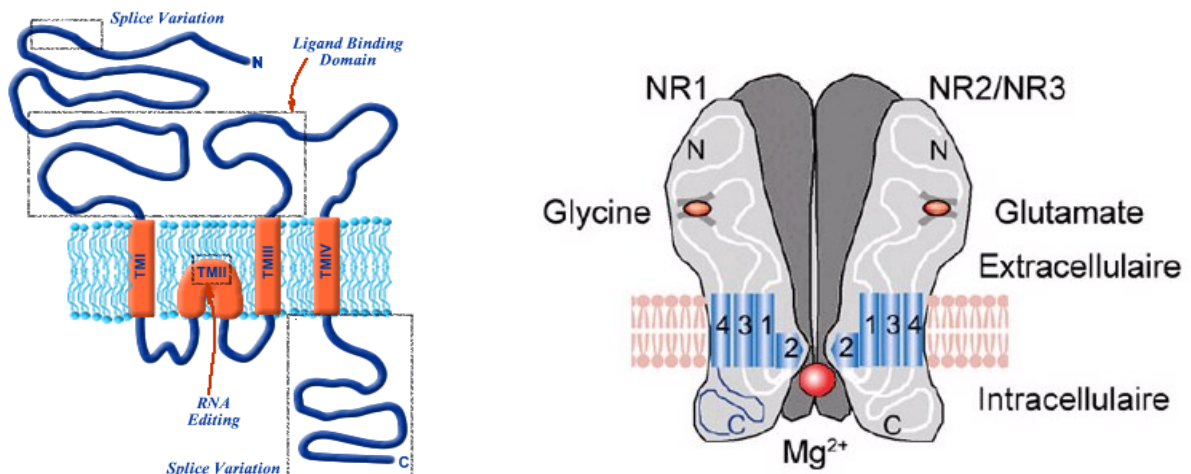


Figure 10 : Structure des récepteurs NMDA

Les sous-unités NR1 sont présentes dans tous les récepteurs NMDA fonctionnels. Ils possèdent le site d'ancrage de la glycine et participent à la diversité des récepteurs du fait de huit différents isoformes générés par trois sites d'épissage alternatif, chacune des isoformes ayant des caractéristiques électrophysiologiques et pharmacologiques spécifiques incluant l'inhibition par les protons, la potentiation par des polyamines ou l'inhibition par le zinc.

Les sous-unités NR2 possèdent le site de liaison au glutamate et spécifient les propriétés électrophysiologiques et pharmacologiques du récepteur, telles que leur affinité au glutamate, leur modulation par la glycine, l'inhibition par le Mg^{2+} ou la cinétique du canal. Dans la mesure où les différentes sous-unités NR2 montrent aussi des patterns d'expression spatiaux et temporels spécifiques il a été suggéré que ces sous-unités pourraient jouer un rôle unique dans le développement cérébral (Cull-Candy & Leszkiewicz 2004). Parmi les sous-unités NR2, NR2A et NR2B en particulier sont exprimés dans toutes les aires du cerveau, particulièrement celles impliquées dans

l'apprentissage et la mémoire comme l'hippocampe ou le cortex. Au cours du développement, il semble qu'il y ait un remplacement des récepteurs contenant des sous-unités NR2B par des récepteurs contenant des sous-unités NR2A, ce mécanisme passe probablement par une perte de localisation synaptique des sous-unités NR2B, suite au changement de composition des molécules d'échafaudage présentes d'une part à la synapse et d'autre part dans les zones extrasynaptiques, le niveau d'expression de NR2B étant constant.

Bien que, typiquement, les récepteurs NMDA requièrent à la fois la présence de glutamate et de glycine pour leur activation, les sous-unités NR3 ne possèdent pas de sites de fixation au glutamate mais possèdent, tout comme les sous-unités NR1, un site de liaison pour la glycine. Ainsi les récepteurs composés uniquement de sous-unités NR1 et NR3 sont activés par la glycine seule. En revanche lorsque les sous-unités NR3 sont coexprimées avec les sous-unités NR2 et NR1 elles peuvent agir en tant que modulateur négatif et réduire la conductance et la perméabilité au Ca^{2+} (Das et al 1998).

L'assemblage des récepteurs sous forme tétramérique est effectué dans le réticulum endoplasmique et les récepteurs fonctionnels sont transportés sous forme vésiculaire jusqu'au site synaptique par des mécanismes actifs impliquant des protéines motrice, cependant des mécanismes de sortie du réticulum de sous-unités unique est possible. A ce jour le seul transporteur moléculaire identifié est la kinésine 17. Cette protéine fait partie de la famille des kinésines qui sont des moteurs moléculaires ancrés aux microtubules et impliqués dans le transport de différents composants membranaires, complexes protéiques et ARNm (Hirokawa et al 2010). Toutes les kinésines possèdent un domaine

globulaire moteur relativement similaire, en revanche les autres domaines sont divergents et permettent l'ancrage à différentes molécules d'ancrage (molécules cargo) permettant la diversité des composants transportés (Hirokawa & Takemura 2005). La kinésine 17, en particulier, joue un rôle crucial dans le transport vésiculaire dendritique des récepteurs NMDA contenant la sous-unité NR2B (Setou et al 2000). Elle est aussi impliquée dans le transport de la sous-unité GluR5 des récepteurs kainate (Kayadjanian et al 2007) et dans le transport du canal potassium Kv4.2 (Covarrubias et al 2008).

En plus de la complexité générée par la composition en sous-unités du récepteur la diversité est augmentée du fait de la diversité des protéines de la densité postsynaptique auxquels les récepteurs peuvent se lier mais aussi du fait que les sous-unités contiennent de nombreux sites régulateurs capable de lier le zinc, les protons ou des polyamines (spermine, spermidine) par exemple.

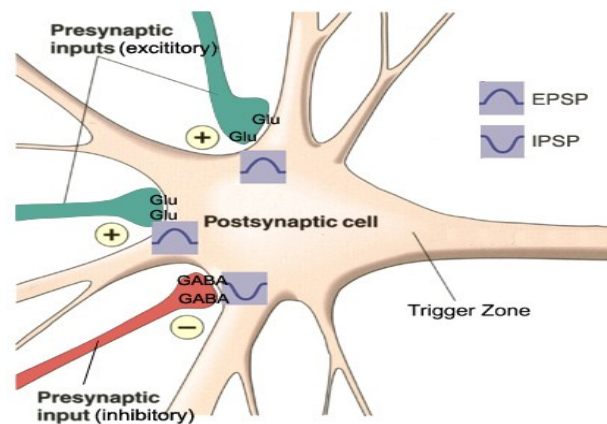


Figure 11 : Systèmes GABA et glutamatergiques.

Tiré de (Occhipinti et al., *Journal of Cerebral Blood Flow & Metabolism*, 2010)

(Occhipinti et al 2010).

III-NMDA et schizophrénie

A-Introduction à l'hypothèse glutamatergique

Un dysfonctionnement dans la transmission glutamatergique en général et du système NMDA en particulier est suspecté dans l'étiologie de la schizophrénie. Cette hypothèse a émergé dans les années 1980, du fait d'observations cliniques (exposition à substances psychotropes d'action glutamatergique) et scientifiques convergentes et a permis de combler certaines des insuffisances de la théorie dopaminergique, théorie la plus explorée jusqu'alors. L'hypothèse glutamatergique propose qu'il y aurait une convergence de facteurs de risque génétique liés aux récepteurs NMDA qui pourraient expliquer les anomalies d'activité dopaminergique observées chez les patients par l'intermédiaire de la modulation des récepteurs NMDA sur les neurones dopaminergiques. Il a été suggéré que le glutamate pourrait être impliqué dans l'étiologie de la maladie de par son rôle dans l'apoptose au cours du développement (Glantz et al 2006; Jarskog et al 2005). Cette théorie repose notamment sur les observations d'anomalies cérébrales structurelles mais l'absence de gliose réactionnelle à la mort cellulaire, typique des maladies neurodégénératives comme les maladies d'Alzheimer et de Parkinson, n'est pas en faveur d'une implication neurodégénérative dans la schizophrénie et plaide plutôt pour la théorie neurodéveloppementale (Weinberger & McClure 2002).

La première observation significative d'une anomalie glutamatergique dans la schizophrénie provient de l'observation d'une diminution des concentrations de glutamate dans le LCR de patients souffrant de schizophrénie (Kim et al 1980) et, sur des

critères pharmacologiques, physiopathologiques et génétiques, la recherche s'est orientée principalement sur le cas des récepteurs NMDA.

B-Arguments pharmacologiques

Un des arguments les plus forts d'une implication des récepteurs NMDA dans la schizophrénie provient de l'observation des effets comportementaux des antagonistes NMDA (PCP, kétamine) à la fois chez le sujet sain et chez l'animal. Que ce soit en administration chronique ou aigue, les antagonistes NMDA sont capables d'induire chez le sujet sain des effets comportementaux, cognitifs et physiologiques proches des symptômes observés chez les patients (Amitai & Markou 2010; Javitt & Zukin 1991; Krystal et al 2002; Krystal et al 1994). Chez les patients, ils sont capables d'exacerber les symptômes ou de faire réapparaître d'anciens symptômes (Javitt & Zukin 1991; Krystal et al 1994).

Ainsi, ces antagonistes semblent mieux « mimer » la symptomatologie de la maladie que les agonistes dopaminergiques, qui induisent exclusivement des symptômes positifs. Ces observations suggèrent fortement une diminution de la fonction des récepteurs NMDA dans la physiopathologie de la maladie et ont incité à explorer des modèles pharmacologiques de la maladie chez l'animal. Chez les patients, des essais thérapeutiques avec des molécules activant directement ou indirectement le site de modulation de la glycine ont montré une certaine efficacité significative sur la diminution des symptômes négatifs, sur l'augmentation des capacités cognitives et, dans un certain nombre de cas, dans la réduction des symptômes positifs. L'amélioration reste modérée et le bénéfice thérapeutique reste à démontrer (Tuominen et al 2006).

C-Arguments physiopathologiques

De très nombreuses études ont mis en évidence soit en post-mortem soit *in vivo* grâce aux travaux de spectrométrie des anomalies d'expression des récepteurs NMDA chez les patients atteints de schizophrénie, y compris chez les patients n'ayant jamais reçu d'antipsychotiques. Ces anomalies peuvent concerner le niveau d'expression de ces sous-unités (Meador-Woodruff & Healy ; Pilowsky et al ; Watis et al) mais aussi la composition spécifique en sous-unités du récepteur (Akbarian et al 1996; Gao et al 2000; Ibrahim et al 2000; Kristiansen et al 2007). Enfin, un nombre croissant d'études semble indiquer un lien entre encéphalites à anticorps anti-NMDA et schizophrénie (Dalmau et al 2008; Dalmau et al 2011; Vincent & Bien 2008).

D-Arguments génétiques

Certaines études génétiques sont en faveur d'une association entre certains variants fréquents de GRIN2A et GRIN2B et schizophrénie. En particulier pour GRIN2B, il a été montré un effet fonctionnel des SNP associés à la maladie avec une modification de l'expression du gène. Cependant la plupart de ces résultats n'ont pas été répliqués et que l'effet de chacun des variants est sans doute relativement modéré.

	Nbre de SNP testés	Etudes
NR1	12	4 positives (Toutes pour rs11146020 dans le promoteur) 9 négatives (Dont 4 pour rs11146020)
NR2A	6	1 positive et 2 avec tendance (Asian) Toutes pour un stretch de GT dans le promoteur 2 négatives (1 pour le stretch de GT)
NR2B	31	5 positives (SNP dans les 5' et 3'UTR, Modification des niveaux d'expression) 6 négatives (3 pour les SNP en UTR)
NR2C	3	1 négative
NR2D	9	1 positive 1 négative
NR3A	2	1 négative

Figure 12 : Etudes d'association et sous-unités du récepteur NMDA

Bien qu'il n'y ait pas de preuves génétiques directes de l'implication des récepteurs NMDA dans la schizophrénie, un des modèles les plus convaincants de schizophrénie chez l'animal est celui de la perte de fonction de NR1 dans environ la moitié des interneurons corticaux, au cours du développement prénatal. Dans ce modèle, les souris présentent des déficits apparaissant à l'adolescence en accord avec les déficits attendus dans un modèle animal de schizophrénie, à savoir des déficits de la mémoire de travail, du prepulse inhibiton, de l'anhédonie ou des comportements anxieux (Belforte et al 2010). D'autres modèles basés sur des modifications génétiques de l'expression des récepteurs NMDA ont aussi reçu des résultats convaincants (Cheli et al 2006; Inada et al 2003; Mohn et al 1999).

Par ailleurs, de nombreux gènes impliqués dans les études génétiques (y compris ceux présentés dans ce travail) jouent un rôle dans la mise en place de la synapse glutamatergique, l'activité des récepteurs glutamatergiques et des récepteurs NMDA en particulier (Banerjee et al 2010; Hahn et al 2006; Harrison & Weinberger 2005; Lau & Zukin 2007; Pitcher et al 2011; Ross et al 2006) (voir **figure 13**). Cela inclut par exemple Neuregulin 1 (NRG1) et le récepteur erbB4. Ainsi il a été suggéré récemment que les anomalies associées à cette voie sont dues à un mécanisme de régulation de l'activité des récepteurs NMDA par l'intermédiaire de leur phosphorylation par la tyrosine kinase Src (Hahn et al 2006; Pitcher et al 2011). Aussi cela inclut la neurexin et les neuroligin qui permettent la mise en place de la densité post synaptique et le recrutement des récepteurs glutamatergiques et qui ont été impliqué dans les études de CNV (Kirov et al 2008; Rujescu et al 2009; Walsh et al 2008) ou SHANK3, associé à l'autisme (Durand et al 2007; Gauthier et al 2009) et qui a récemment été associé à la schizophrénie (Gauthier et al 2010).

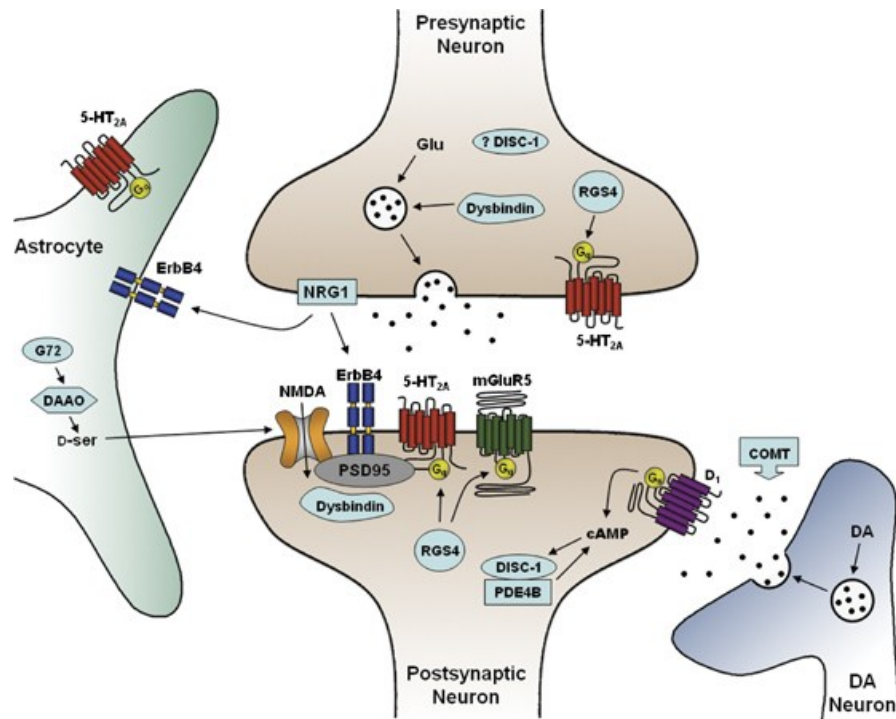


Figure 13 : Synapse glutamatergique et schizophrénie

(Tiré de Gray et Roth(Gray & Roth 2007))

A ce jour, deux études ont eu pour objectif de séquencer les gènes des sous-unités NMDA dans des maladies neurodéveloppementales. La première étude (Williams et al 2002) chez des patients atteints de schizophrénie s'est focalisée sur les gènes *GRIN1* et *GRIN2* mais n'a pas permis l'identification de variants ayant potentiellement un effet fonctionnel délétère et seulement un polymorphisme non synonyme. De façon intéressante, la deuxième étude, plus récente (Endele et al 2010), a permis l'identification de mutations rares et d'anomalies chromosomiques touchant ces gènes chez des patients atteints de déficience mentale.

Les nouvelles technologies permettent aujourd'hui de rechercher avec une méthodologie à haut débit les variants rares. Dans la mesure où les travaux épidémiologiques et en

imagerie confirment l'hypothèse qu'une anomalie du neurodéveloppement peut mener à un trouble schizophrénique et que dans d'autres troubles développementaux comme l'autisme, des mutations rares à effet dominant ont été montrées, modifiant le fonctionnement de la synapse glutamatergique, il est pertinent de rechercher de telles mutations chez des sujets atteints de schizophrénie, d'autisme ou de déficience intellectuelle.

Notre travail s'est concentré sur l'identification de variants rares à effet dominant dans les gènes des récepteurs NMDA et dans le gène de la kinésine kif17, protéine impliquée dans le transport des sous unités NR2B dans la schizophrénie.

Chapitre 2 : travail expérimental

Contexte : le projet « synapse to disease » : S2D

Le projet a vu le jour en 2006 à la suite d'un financement initial octroyé par Génome Canada et Génome Québec pour l'étude de l'autisme et de la schizophrénie. Il met à contribution un grand nombre de partenaires et avait pour objectif initial de séquencer 1000 gènes synaptiques dans des cohortes de patients atteints d'autisme, de schizophrénie, ainsi que leurs parents ; les gènes de la fonction synaptique étant surreprésentés à la fois par les études d'association et par les études récentes de CNV (Glessner et al 2010; Guilmatre et al 2009). La stratégie employée a été d'inclure des patients n'ayant pas d'histoire familiale de troubles psychiatriques afin de maximiser les probabilités d'identifier des mutations *de novo*, ce type de mutations étant suggéré par plusieurs études. Dans ce contexte, l'implication causale d'une mutation ponctuelle observée chez le sujet atteint est rendue d'autant plus plausible que (i) ses conséquences fonctionnelles sont avérées (protéine tronquée perdant un domaine fonctionnel par exemple) (ii) la mutation est absente chez les parents sains et chez les contrôles.

Lorsque je suis arrivé fin janvier 2008 dans le laboratoire du Dr Rouleau, dans le cadre de ma cotutelle de thèse, les étapes initiales du projet, à savoir la sélection des gènes synaptiques et des individus, était terminée et le séquençage était déjà amorcé.

La recherche des gènes synaptiques s'est effectuée sur la base de recherche bibliographique sur différentes bases de données. La sélection a ensuite été exécutée par attribution d'un score pour chacun des gènes. Les critères pour l'attribution des scores de sélection incluaient des éléments qui pourraient laisser penser à une susceptibilité d'impliquer ces gènes dans l'étiologie de ces maladies. Cela impliquait des critères

fonctionnels (Rôle dans la plasticité synaptique, l'apprentissage et la mémoire par exemple), d'expression tissulaire (expression spécifique dans le cerveau ou non, étude par microarrays), génétiques (Études d'association, CNV, rôle dans d'autres maladies (OMIM)) et protéomiques.

Les gènes présentant un maximum d'arguments étant ceux sélectionnés pour le séquençage.

Pour la sélection des individus la stratégie étant d'essayer d'identifier des mutations *de novo* les patients sélectionnés en priorité devaient, de préférence, ne pas avoir d'histoire familiale de troubles psychiatriques et avoir un âge des parents élevés. La condition *sine qua none* pour identifier des mutations *de novo* étant d'avoir accès à l'ADN sanguin parental.

Le séquençage a été effectué par la technique classique de Sanger et concernait les régions exoniques des gènes sélectionnés. Les cohortes de patients prévues pour l'étape initiale du projet incluaient 142 patients atteints de troubles du spectre autistique et 143 patients atteints de schizophrénie.

Les parents de ces individus et une cohorte de contrôle (au départ 270 contrôles provenant d'une vaste étude longitudinale de jumeaux) étaient disponibles pour l'analyse des variants les plus pertinents (mutations tronquantes principalement)

Les étapes suivantes concernaient la détection. La validation et l'analyse des variants.

La détection des mutations a été faite à l'aide de logiciels d'alignement tels que polyphen et polyscan. Chacun des variants détecté par ces logiciels était ensuite confirmé, d'abord par analyse manuelle des séquences à l'aide de Mutation Surveyor puis validé par reséquençage chez le patient et chez ses parents. De même, l'effet fonctionnel de chacun

des variants a été prédit *in silico* à l'aide de logiciels de prédictions. Ces analyses incluent par exemple l'effet fonctionnel potentiel, de la modification d'un acide aminé, de la création ou de la délétion d'un site d'épissage.

Au final une dernière étape de validation biologique était prévue pour la validation de l'effet fonctionnel des variants les plus pertinents, plusieurs modèles étant disponibles à savoir des modèles cellulaires (Cellules HEK, cultures neuronales) et des modèles animaux (*C.elegans*, drosophile, zebrafish, souris).

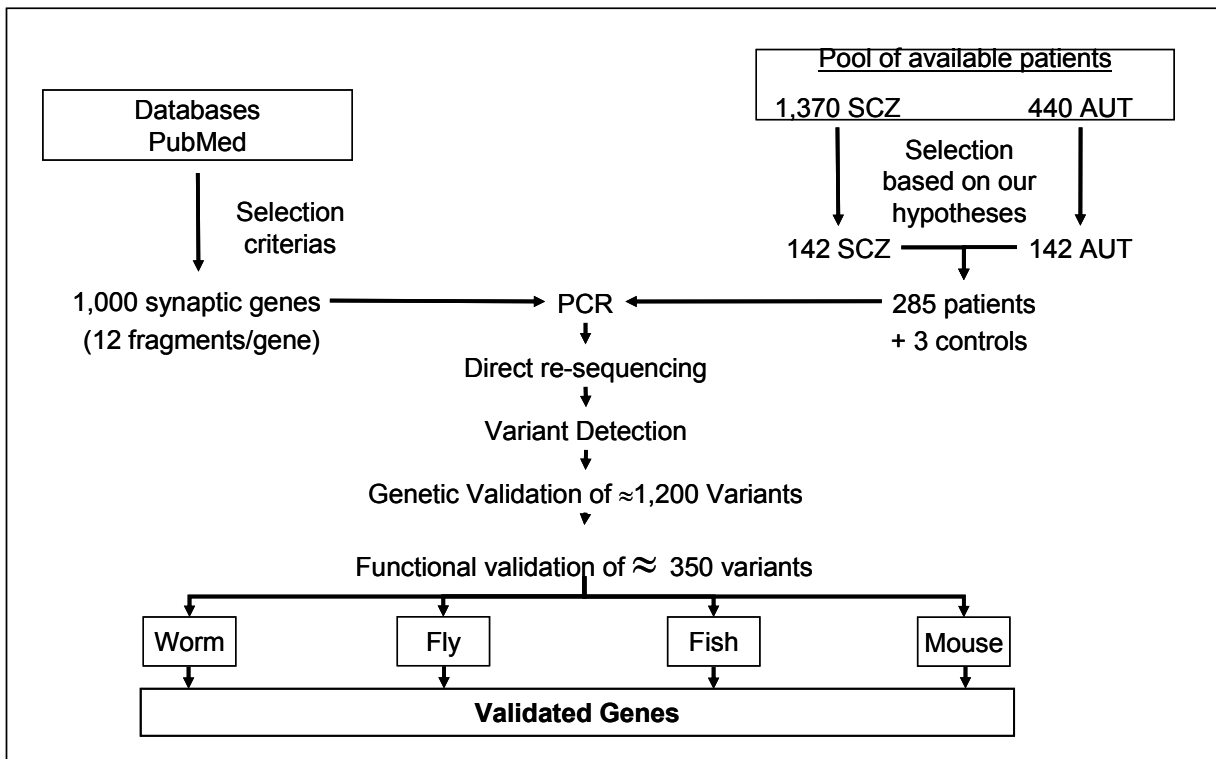


Figure 14 : présentation du projet S2D

Article 1 : De Novo Truncating Mutation in KIF17 Associated with Schizophrenia.

Introduction

Suivant la ligne stratégique du projet S2D, et en accord avec l'idée qu'une fraction des cas de schizophrénie pourrait résulter de variants *de novo*, le premier travail s'est focalisé sur l'analyse des variants rares observés après le séquençage du gène de la kinésine 17.

Cette protéine est impliquée notamment dans le transport vésiculaire dendritique des récepteurs NMDA contenant la sous-unité NR2B.

Nous avons séquencé l'ensemble du gène *KIF17* codant la kinésine 17 chez 188 patients atteints de schizophrénie mais n'ayant pas d'histoire familiale de la maladie, 142 patients atteints de troubles du spectre autistique et 95 patients atteints de déficience intellectuelle. Enfin nous avons séquencé ce gène chez 568 individus contrôles issus d'une part d'une population générale québécoise, d'autre part d'une population de contrôles sains n'ayant pas de troubles psychiatriques.

Nous avons déterminé la transmission de tous les variants uniques (apparaissant une seule fois et non reportée dans les études précédentes) par séquençage du fragment concerné chez les parents des individus (parents disponibles pour 285 des 568 individus contrôles).

Nous rapportons dans cet article l'identification d'une mutation *de novo* non-sens chez un patient sporadique atteint de schizophrénie.

De Novo Truncating Mutation in Kinesin 17 Associated with Schizophrenia

Julien Tarabeux, Nathalie Champagne, Edna Brustein, Fadi F. Hamdan, Julie Gauthier, Mathieu Lapointe, Claudia Maïos, Amélie Piton, Dan Spiegelman, Édouard Henrion, Synapse to Disease Team, Bruno Millet, Judith L. Rapoport, Lynn E. DeLisi, Ridha Joobier, Ferid Fathalli, Éric Fombonne, Laurent Mottron, Nadine Forget-Dubois, Michel Boivin, Jacques L. Michaud, Ronald G. Lafrenière, Pierre Drapeau, Marie-Odile Krebs, and Guy A. Rouleau

Background: Schizophrenia (SCZ) is one of the most disabling psychiatric disorders. It is thought to be due to a complex interplay between polygenic and various environmental risk factors, although recent reports on genomic copy number variations suggest that a fraction of the cases could result from variably penetrant de novo variants. The gene encoding the synaptic motor protein kinesin 17 (KIF17) involved in glutamatergic synapse is a candidate gene for SCZ.

Methods: As part of our Synapse to Disease project, we resequenced KIF17 in a cohort of individuals with sporadic SCZ (188 subjects). Additional populations included autism spectrum disorder (142 subjects), nonsyndromic mental retardation (95 subjects), and control subjects (568 subjects). Functional validation of the human mutation was done in developing zebrafish.

Results: Here we report the identification of a de novo nonsense truncating mutation in one patient with SCZ, in kinesin 17, a synaptic motor protein. No de novo or truncating KIF17 mutations were found in the additional samples. We further validated the pathogenic nature of this mutation by knocking down its expression in zebrafish embryos, which resulted in a developmental defect.

Conclusions: Together our findings suggest that disruption of KIF17, although rare, could result in a schizophrenia phenotype and emphasize the possible involvement of rare de novo mutations in this disorder.

Key Words: De novo, KIF17, kinesin, NMDA, schizophrenia, synapse

Schizophrenia (SCZ) is a prevalent and severe psychiatric illness that usually begins in adolescence or early adulthood and has severe consequences on the health of affected individuals and important economical burden on the

society (for general review, see [1]). Prior to disease onset, a substantial number of patients present mild motor, social, or cognitive dysfunction supporting the notion that SCZ is a neurodevelopmental disorder, a hypothesis further supported by epidemiological, brain imaging, and postmortem studies (2,3).

High heritability (80%) has been demonstrated in SCZ and the prevailing view suggests that a combination of multiple common variants and various environmental factors could account for its pathophysiology. Nevertheless, despite genomewide meta-analysis including 32 linkage studies (3255 pedigrees and 7413 cases) (4) and genomewide association studies (5–8) in extensive samples of patients and control subjects (more than 30,000 individuals), the unambiguous identification of such common susceptibility variants is yet to be achieved. Clinical and genetic heterogeneity of SCZ could actually overwhelm the potency of these strategies. In addition to the polygenic hypothesis, there is growing evidence that at least some cases could be associated with rare variably penetrant de novo mutations. This alternative, but noncontradictory, hypothesis has found recent support from the findings of a relatively high frequency of rare de novo copy number variations in patients with SCZ (6,9,10) similar to what had been reported in other neurodevelopmental disorders such as autism and mental retardation. It is, however, still debated whether those reported rare copy number variations are associated with SCZ itself or rather, less specifically, with mental retardation or autism secondarily associated with comorbid psychotic symptoms. When available in these reports, clinical descriptions so far have indeed underlined the presence of either autistic-like symptoms, early onset of SCZ, and/or mild mental retardation in the patients.

The biological mechanisms underlying SCZ remain largely unknown, although there is increasing evidence that defective synaptic function could account for both disrupted development and psychiatric and cognitive symptoms in patients (11–13). We

From the Center of Excellence in Neuromics of Université de Montréal, Centre hospitalier de l'Université de Montréal Research Center, and the Department of Medicine (JT, JG, AP, DS, EH, RGL, GAR, Synapse to Disease Team), University of Montréal, Montreal, Quebec, Canada; Institut national de la santé et de la recherche médicale Laboratory of Pathophysiology of Psychiatric Diseases, Center of Psychiatry and Neurosciences (JT, BM, M-OK), and University Paris Descartes, Centre d'Évaluation et de Recherche Clinique (JT, BM, M-OK), Sainte-Anne Hospital, Paris, France; Department of Pathology and Cell Biology and Le Groupe de Recherche sur le Système Nerveux Central (NC, EB, ML, CM, PD, Synapse to Disease Team) and Research Center, Centre hospitalier universitaire Sainte-Justine, and Department of Medical Genetics (FFH, JLM), Université de Montréal, Montreal, Quebec, Canada; Service Hospitalo-Universitaire de Psychiatrie, Centre hospitalier Guillaume Rénier, Rennes, France; Child Psychiatry Branch (JLR), National Institute of Mental Health, Bethesda, Maryland; The Boston Veterans Affairs Healthcare System (LED), Harvard Medical School, Brockton, Massachusetts; Douglas Mental Health University Institute, Department of Psychiatry (RJ, FF), and Department of Psychiatry, Montreal Children's Hospital (EF), McGill University, Montreal, Quebec, Canada; Centre de Recherche Fernand-Séguin (LM), Hôpital Rivière-des-Prairies, Université de Montréal, Montreal, Quebec, Canada; and Canada Research Chair on Child Social Development, Research Unit on Children's Psychosocial Maladjustment (NF-D, MB), Laval University, Quebec City, Quebec, Canada.

Address correspondence to Guy A. Rouleau, M.D., Ph.D., CHUM Research Centre, 2099 Alexandre-De-Seve Street, Rm Y-3633, Montreal, Quebec, Canada H2L 2W5;

Received Dec 10, 2009; revised Apr 9, 2010; accepted Apr 9, 2010.

Discussion

Ce papier est le premier à rapporter le cas d'une mutation dans *KIF17* chez un cas de schizophrénie. Le tableau présenté est typique, sans déficience intellectuelle marquée et avec un âge de début pas particulièrement précoce, si on se réfère à l'âge d'apparition des troubles délirants.

Le caractère fonctionnel de la mutation a été vérifié contrôlant en Western Blot que le gène muté mène à la synthèse d'une protéine tronquée qui perd 454 acides aminés. Nous avons confirmé fonctionnellement *in vivo* l'effet de notre mutation dans un modèle utilisant le zébrafish et nous avons pu montrer que la protéine résultant du gène muté n'est pas fonctionnelle, avec un phénotype plutôt sévère modifiant le développement du zébrafish.

Aucune autre mutation *de novo* ou tronquante du gène *KIF17* n'a été identifiée, ni dans la cohorte initiale de patients sporadiques atteints de schizophrénie (n=188), ni dans les cohortes additionnelles incluant 142 patients atteints de troubles du spectre autistique, 95 patients souffrant de déficit intellectuel et 568 individus contrôles suggérant que cette mutation est spécifique de la maladie.

L'absence de mutation chez les parents est un argument supplémentaire en faveur du rôle de la mutation dans le tableau psychiatrique observé, dans la mesure où les deux parents ont été évalués comme n'ayant pas de troubles psychiatriques.

La kinésine 17 est un moteur moléculaire impliqué dans le transport vésiculaire le long des microtubules de différents récepteurs membranaires. Elle est impliquée dans le transport vésiculaire dendritique des récepteurs NMDA contenant la sous-unité NR2B (Setou et al 2000), dans le transport de la sous-unité GluR5 des récepteurs kainate (Kayadjanian et al 2007) et dans le transport du canal potassium Kv4.2 (Covarrubias et al 2008).

In vitro, elle joue un rôle dans le niveau d'expression synaptique des sous-unités NR2B (Guillaud et al 2003). Aussi, elle a été impliquée dans des processus de mémoire et d'apprentissage (Wong et al 2002; Yin et al 2011) et, *in vivo* chez la souris knockout pour *kif17* il a été montré que les déficits observés sont principalement dus à des variations dans l'expression synaptique des récepteurs NMDA, suggérant des mécanismes de transport redondants pour les autres molécules transportés par cette kinésine mais pas pour les récepteurs NMDA contenant la sous-unité NR2B (Yin et al 2011).

Au final, pris dans leur globalité, nos résultats suggèrent que la perte de fonction de *KIF17* est associée à la schizophrénie, impliquant probablement des anomalies du système NMDA, et accentuent l'implication probable, de mutations rares *de novo* dans l'étiologie de cette maladie.

***Article 2 : Rare mutations in N-méthyl-D-aspartate (NMDA) glutamate receptors
in autism spectrum disorders and schizophrenia.***

En préparation,

Introduction

L'implication du système glutamatergique, et du système NMDA en particulier, dans l'étiologie de la schizophrénie est supportée par les études pharmacologiques, génétiques, d'expression ainsi que par l'article précédent.

Dans cet article, nous avons exploré les mutations rares et *de novo* dans les sous-unités des récepteurs NMDA. Nous avons séquencé tous les exons des sept gènes codant pour des sous-unités NMDA chez 571 patients, incluant 429 patients atteints de schizophrénie et 142 patients atteints de troubles du spectre autistique. De plus, nous avons séquencé tous les gènes (sauf *GRIN3B* et *GRIN1*) dans une population de 283 volontaires « sains ». Nous avons séquencé des cohortes additionnelles incluant 286 patients atteints de troubles du spectre autistique pour *GRIN2B*, *GRIN2C* et *GRIN3B* et 285 individus provenant de la population générale pour *GRIN2C* et *GRIN3A*. Pour *GRIN1* nous avons séquencé uniquement cette deuxième cohorte d'individus de la population générale.

Nous avons essayé de déterminer la transmission de tous les variants uniques en séquençant les parents de l'individu porteur de la mutation (292 patients atteints de schizophrénie, 404 atteints d'autisme et 285 individus de la population générale avec les 2 parents disponibles).

Nous rapportons dans ce papier l'identification de plusieurs mutations potentiellement causales et/ou *de novo*. En particulier des mutations *de novo* pouvant être associée à la schizophrénie et aux troubles du spectre autistique.

Rare mutations in N-methyl-D-aspartate (NMDA) glutamate receptors in autism spectrum disorders and schizophrenia.

Julien Tarabeux^{1,2,3}, Oussama Kébir^{2,3}, Julie Gauthier¹, Fadi F. Hamdan⁴, Lan Xiong¹, Amélie Piton¹, Dan Spiegelman¹, Édouard Henrion¹, S2D team^{1,5}, Bruno Millet², Ferid Fathalli⁶, Ridha Joobar⁶, Judith L. Rapoport⁷, Lynn E. DeLisi⁸, Éric Fombonne¹⁰, Laurent Motttron¹¹, Nadine Forget-Dubois¹², Michel Boivin^{12,13}, Jacques L. Michaud⁴, Pierre Drapeau⁵, Ronald G. Lafrenière¹, Guy A. Rouleau^{1*CA}, Marie-Odile Krebs^{2,3*CA}.

¹Center of Excellence in Neuroscience of the Université de Montréal (CENUM), Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), and Department of Medicine, University of Montreal, Montreal, Quebec, Canada, H2L 4M1. ²INSERM Laboratory of Pathophysiology of Psychiatric Diseases, Center of Psychiatry and Neurosciences, U894, Sainte-Anne Hospital, Paris, France. ³University Paris Descartes, Faculty of Medicine Paris Descartes, Centre d'Evaluation et de Recherche Clinique, Sainte-Anne Hospital, Paris, France. ⁴ Centre of Excellence in Neuroscience of Université de Montréal (CENUM), Centre de Recherche du CHU Sainte-Justine, Montréal, Quebec, H3T1C5, Canada. ⁵Department of Pathology and Cell Biology and Le Groupe de Recherche sur le Système Nerveux Central, Université de Montréal, Pavillon Roger-Gaudry, Montréal, (QC), Canada. ⁶Douglas Mental Health University Institute, Department of Psychiatry, McGill University, Pavillon Frank B. Common, Verdun, Montreal, (QC), Canada. ⁷Child Psychiatry Branch, National Institute of Mental Health, Bethesda, MD, USA. ⁸Harvard Medical School, The Boston VA Healthcare System, Brockton, MA, USA. ¹⁰Department of Psychiatry,

Montreal Children's Hospital, McGill University, Montreal, (QC), Canada. ¹¹Hôpital Rivière-des-Prairies, Centre d'excellence en Trouble envahissants du développement de Université de Montréal (CETEDUM), Montréal, (QC), Canada. ¹²Research Unit on Children's Psychosocial Maladjustment, Laval University, Quebec City, (QC), Canada . ¹³Canada Research Chair on Child Social development.

* the last co-authors equally contributed to the work

Corresponding authors:

Guy ROULEAU : Guy A. Rouleau, MD, PhD, FRCPC, OQ, CHUM Research Centre, 2099 Alexandre-De-Seve Street, Rm Y-3633, Montreal, Quebec, Canada H2L 2W5, Tel: 514-890-8000 x 24699, Fax: 514-412-7602

Marie-Odile KREBS : Laboratoire de Physiopathologie des Maladies Psychiatriques, Centre de Psychiatrie et Neurosciences, UMR 894 INSERM- Université Paris Descartes, Faculté de Médecine Paris Descartes, Service Hospitalo Universitaire, Hôpital Sainte-Anne, 7 rue Cabanis, 75014 Paris, France.

Pharmacological, genetic and expression studies implicate N-methyl-D-aspartate (NMDA) receptor hypofunction in schizophrenia. Similarly, several lines of evidence suggest that autism spectrum disorders (ASD) could be due to an imbalance between excitatory and inhibitory neurotransmission. As part of a project aimed at exploring rare and/or *de novo* mutations in neurodevelopmental disorders, we have sequenced the seven genes encoding for NMDA receptor subunits (NMDARs) in a large cohort of individuals affected with SCZ or ASD (n=429 and 428, respectively), parents of these subjects and controls (n=568). Here, we identified two *de novo* mutations in patients with sporadic SCZ in *GRIN2A* and one *de novo* mutation in *GRIN2B* in a patient with ASD. Truncating mutations in *GRIN2C*, *GRIN3A* and *GRIN3B* were identified in both subjects and controls, but no truncating mutations were found in the *GRIN1*, *GRIN2A*, *GRIN2B* and *GRIN2D* genes both in patients and controls, suggesting that these subunits are critical for neurodevelopment. The present results support the hypothesis that rare *de novo* mutations in *GRIN2A* or *GRIN2B* can be associated with cases of sporadic SCZ or ASD, just as it has recently been described for the related neurodevelopmental disease intellectual disability. The influence of genetic variants appears different depending on NMDAR subunits. Functional compensation could occur to counteract the loss of one allele in *GRIN2C* and *GRIN3* family genes while *GRIN1*, *GRIN2A*, *GRIN2B* and *GRIN2D* appear instrumental to normal brain development and function.

Introduction

Glutamate is the major excitatory neurotransmitter of the central nervous system (CNS) and is implicated in many basic neuronal functions and CNS processes, in particular learning, memory and synaptic plasticity. N-methyl-D-aspartate receptors (NMDAR) are voltage dependant ionotropic glutamate receptors and are tetramers containing two NR1 subunits in association with two NR2 (A, B, C or D) and/or NR3 (A or B) subunits. Many of the physiological and pharmacological properties of the NMDAR depend on the specific NR2 and NR3 subunits composition.¹ Behavioral tasks that involve NMDAR include associative learning, working memory, behavioral flexibility or attention. Thus, NMDAR signalling could be a point of convergence to explain characteristic patterns of symptoms and neurocognitive deficits observed in neurodevelopmental disorders.

In schizophrenia, dysregulation of brain glutamate transmission through NMDAR has been proposed as an etiologic factor. Indeed, NMDAR antagonists (e.g. ketamine, phencyclidine) produce behavioral and cognitive deficits in normal subjects that closely mimic schizophrenia²⁻⁴, and patients with anti-NR1 encephalitis display many schizophrenia-like symptoms and/or loss of memory.⁵⁻⁷ Mutant mice lacking NR1 subunit in about half of cortical interneurons early in post-natal development exhibit anxiety-like behaviours and working memory deficits, emerging at adolescence.⁸ In addition, NMDAR density and NMDAR subunit composition are reported abnormal in post-mortem brains of patients with schizophrenia.⁹⁻¹⁵

Schizophrenia has been associated with genetic variants in some of the NMDAR subunit (NMDARs) genes including *GRIN2A*, *GRIN2B* and *GRIN1*, although contradictory results were reported (see the SCZgene database : <http://www.szgene.org/>).¹⁶ In addition, several

genes that have been associated with increased risk for schizophrenia, including *NRG1-erb4*, *serine racemase*, *DAO*, *G72* or *dysbindin-1*, are implicated in the modulation of NMDAR activity.¹⁷⁻¹⁹ Lastly, the only consistent report from independent genome wide association studies in schizophrenia is the association with genetic markers in the major histocompatibility complex region (MHC)²⁰⁻²² and recent evidence suggests that MHC is implicated in neuronal plasticity²³ and glutamatergic receptor modulation.²⁴

Dysfunction of NMDA signalling has also been implicated in a wide range of neurological or neurodevelopmental disorders such as intellectual disability (ID)²⁵⁻²⁶ where it appears that specific NMDARs may play different roles in synaptic plasticity and excitotoxicity.²⁷ In autism spectrum disorders (ASD), potentially causative mutations in NMDARs genes have been recently reported²⁸⁻²⁹ and a number of ASD animal models were associated with NMDA abnormalities.³⁰⁻³³ Moreover, disruption in the balance between excitation and inhibition is a commonly proposed disease mechanism for ASD³⁴ and an open-label study with memantine, an antagonist of the NMDAR, significantly improved symptoms.³⁵ Altogether, the literature suggests that NMDAR number and composition play a crucial role in neurodevelopment and cognition.

Recent studies indicate that rare and penetrant *de novo* mutations could account for some cases of schizophrenia. For examples, rare chromosomal abnormalities are more frequent in schizophrenia than in controls, especially rare *de novo* Copy Number Variations (CNVs)^{22, 36-37} similar to what had been reported in other neurodevelopmental disorders such as ASD³⁸ and ID.^{8, 39} In addition, recent reports, including ours, issued from the Synapse to disease (S2D) project, support the implication of *de novo* single nucleotide mutations in ID^{25, 40-42}, ASD^{29, 41, 43} or schizophrenia.⁴⁴⁻⁴⁷ Interestingly, among rare *de novo* variants reported associated with

schizophrenia, two affect genes that may be linked to the glutamatergic synapses : *SHANK3*, encoding a scaffolding protein at the excitatory synapse and *KIF17* encoding a synaptic motor protein, involved in the transport of NR2B to the synapse.⁴⁸

As part of the S2D project, in which we screened 600 synaptic genes, we screened all exons of the seven NMDARs in 571 individuals including 429 patients with schizophrenia and 142 patients with ASD. We also screened all exons of the NMDARs (except for *GRIN3B* and *GRIN1*) in 283 healthy controls. Further we screened additional cohort of ASD patients (n=286) for *GRIN2B*, *GRIN2C* and *GRIN3B* and additional cohort of controls from the general population (n=285) for *GRIN2C* and *GRIN3A*. For *GRIN1* we screened only this general population cohort of controls. Finally, we determined transmission of all unique variants, identified in only one individual, by sequencing parents of the carrier (292 schizophrenia and 404 ASD patients with 2 parents available for screening).

We here report the identification of *de novo* mutations in patients with schizophrenia in the *GRIN2A* gene and in a patient with ASD in the *GRIN2B* gene. Also, we report the absence of truncating mutation in the *GRIN1*, *GRIN2A*, *GRIN2B* and *GRIN2D* genes in patients or controls while we identified truncating mutations in the *GRIN2C*, *GRIN3A* and *GRIN3B* genes in patients and controls.

METHODS

Subject collection

Schizophrenia cohort: Subjects with schizophrenia were selected from over 1000 families ascertained for genetic studies, for which DNA samples were available. The screened cohort included 429 subjects with schizophrenia. We selected in priority patients with no family history of psychiatric disorders (n=188 patient) and patients for which both parents were available (n=292 patients with both parents). In order to ensure accurate diagnoses all individuals were evaluated by experienced investigators (JR, LD, MOK and RJ) using the Diagnostic Interview for Genetic Studies (DIGS) or Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS) and multidimensional neurological, psychological, psychiatric, and pharmacological assessments in the recruiting centres. Family history for psychiatric disorders was also collected using the Family Interview for Genetic Studies (FIGS). All DIGS and FIGS have been reviewed by two or more psychiatrists for a final consensus diagnosis based on DSM-III-R or DSM-IV at each centre. For all probands, specific inclusion criteria for the present study were : (1) The selected proband was definitely affected with schizophrenia only, not schizophreniform psychosis or bipolar disorders with psychosis; (2) In families with multiple affected individuals, we selected the most severe schizophrenia case with early (<18 yrs) or childhood onset (<12 yrs), and/or additional neurodevelopmental problems, such as ID, dyslexia and epilepsy, but not autistic disorder; for the childhood onset schizophrenia (COS) cohort, ID and epilepsy were exclusionary; (3) Family history was well documented. Finally, exclusion criteria included patients with psychotic symptoms mainly caused by alcohol, drug abuse, or other clinical diagnoses including cytogenetic abnormalities. The final panel of the samples of 429 schizophrenia subjects were recruited from different centres, i.e. 1) 28 cases

with childhood-onset schizophrenia and 66 with later onset schizophrenia from Judith L. Rapoport (NIMH, USA). Cases with childhood onset schizophrenia were recruited in USA nationwide and assessed as previously described.⁴⁹ Individuals in this cohort known to carry the VCFS deletion on chromosome 22q11 were excluded. 2) 38 cases from Marie-Odile Krebs (Paris, France).⁵⁰ 3) 84 cases from Lynn E. De Lisi⁵¹ 4) 7 cases of adult onset schizophrenia selected from each of seven large highly consanguineous pedigrees with ≥ 10 affected individuals with schizophrenia and schizoaffective disorders from Pakistan 5) 144 cases from Ridha Joober (Montreal, Canada) including 39 cases from Tunisia⁵² 6) 62 cases from Hungary.⁵³ The detailed description of recruitment and ascertainment strategy, diagnostic instruments and criteria have been reported by each centre in their previous publications.

ASD cohort: The final panel of samples included 428 autistic patients (404 with parental DNA of both parents available). Diagnostic and selection criteria for the ASD subjects are described in detail elsewhere.⁵⁴ We screened 142 autistic patients for the seven NMDAR genes and 286 additional cases were screened for variants in *GRIN2B*, *GRIN2C* and *GRIN3B*. Briefly all subjects were diagnosed using DSM-IV or DSM-III-R criteria, and depending of the recruitment site, Autism Diagnostic Interview-Revised (ADI-R) and the Autism Diagnostic Observation Schedule (ADOS) were used. In addition, the Autism Screening Questionnaire (ASQ) was also completed for all subjects. We excluded patients with an estimated mental age < 18 months, a diagnosis of Rett Syndrome or Childhood Disintegrative Disorder and patients with evidence of any psychiatric and neurological conditions including: birth anoxia, rubella during pregnancy, fragile-X disorder, encephalitis, phenylketonuria, tuberous sclerosis, Tourette and West syndromes.

Control cohorts: Blood samples were obtained from a total of 568 control subjects. The first control series comprised healthy volunteers that were screened and selected using standardized procedures. Only individuals without any neuropsychiatric symptoms or family history of neuropsychiatric problems were included as schizophrenia negative controls. These included 283 adult subjects (131 males; 152 females), 225 of whom were of European origin and 58 were non-Europeans (Tunisian). The second group of controls comprised 285 subjects from a general population control group of mainly French Canadian origin for whom we also had parental DNA (Quebec Newborn Twin Study, QNTS; 150 males, 135 females).⁵⁵ We have systematically sequenced all exons of the seven NMDARs genes (except for *GRIN3B* and *GRIN1*) in the cohort of healthy volunteer controls. When rare truncating mutations were identified, we further screened all exons in additional controls from general population (QNTS). For *GRIN1* we screened only this general population cohort of controls.

DNA preparation. Genomic DNA was extracted from blood using Puregene extraction kit (Gentra System, USA). For certain individuals where blood DNA was limiting, we used DNA isolated from a lymphoblastoid cell line derived from the individual for the screen. In all cases, rare mutations were confirmed using blood derived DNA to rule out variations having arisen during production or growth of the lymphoblastoid cell line.

Paternity testing. Paternity, maternity and unique genetic identification of each individual of all families (subject families and controls) was confirmed using 5 - 14 highly informative unlinked microsatellite markers as described in Gauthier et al.⁴⁶.

Gene screening, variation analysis and bioinformatics. In each proband, we sequenced the coding region and the splice junction of *GRIN1* (chr9 : 140033609-140063207, hg19 assembly), *GRIN2A* (chr16: 9,847,267-10,276,263), *GRIN2B* (chr12: 13,714,410-14,133,022), *GRIN2C* (chr17: 72,838,168-72,856,007), *GRIN2D* (chr19: 48,898,132-48,948,187), *GRIN3A* (chr9:104,331,635-104,500,862). *GRIN3B* (chr19: 1,000,437-1,009,723) was sequenced only in patients. The detailed description of the samples screened for each NMDAR gene are mentioned in table 1. Primers were designed using the ExonPrimer program from the UCSC genome browser (supplementary table 3). PCR products were sequenced at the McGill University and Genome Quebec Innovation Centre in Montreal, Canada (www.genomequebecplatforms.com/mcgill/) on a 3730XL DNA Analyzer System. In each case, variations were confirmed by re-amplifying the fragment and re-sequencing of the samples from the proband and available parents using reverse and forward primers. PolyPhred (version 6.11), PolySCAN (v.3.0) and Mutation Surveyor (version 3.10, Soft Genetics Inc.) were used for mutation detection analysis. The PolyPhen (<http://genetics.bwh.harvard.edu/pph>), SIFT (<http://blocks.fhcrc.org/sift/SIFT.html>), Panther (www.pantherdb.org/tools/csnpscoreform.jsp) and SNAP (<http://www.rostlab.org/services/SNAP/>) programs were used to predict the overall severity of the missense mutations. In all instances, default parameters were used for each program.

Results

We found three *de novo* mutations in patients with sporadic ASD or schizophrenia in *GRIN2A* and *GRIN2B* (table 2). We also identified one truncating mutations in *GRIN2C*, four

truncating mutations in *GRIN3A* and recurrent truncating mutations in *GRIN3B* both in patients and controls (Table 3). By contrast, we found no truncating mutations in patients or in controls in the *GRIN1*, *GRIN2A*, *GRIN2B* and *GRIN2D* genes. All three de novo mutations and truncating mutations in *GRIN2C* and *GRIN3* were not reported neither in the most recent 1000 genome release nor in hapmap.

Two *de novo* mutations were identified in the *GRIN2A* gene. The first in a patient with schizophrenia with no familial history of schizophrenia was a missense (c.2902G>A ; p.A968T) located in the intracellular domain of the protein. It was not predicted as having important functional consequence based on *in silico* analysis (SIFT, PolyPhen and SNAP) (Table 2); nevertheless it modified an amino-acid that is highly conserved in mammals. The second one was a silent mutation: c.3669C>T identified in another patient with sporadic schizophrenia and was not predicted to be damaging by *in silico* analysis. Paternal age was 39 and maternal age was 29 for this patient. These mutations were not found in the remaining 1707 chromosomes that had been sequenced. None of these patients had a mental retardation with 11 and 12 years of education respectively.

A missense mutation (c.2473T>G ; p.L825V) was identified in the *GRIN2B* gene in a patient with ASD that had no family history of neuropsychiatric disorders or ID. The mutation was confirmed by resequencing blood DNA of this patient and his parents. Paternal age was 36 and maternal age was 32 for this patient. This mutation was absent in the remaining 2279 chromosomes and was predicted by *in silico* analysis as relatively damaging as it occurred in the last transmembrane domain that is highly conserved among species.

We further analyzed the prevalence of truncating mutations in our subjects. We did not identify any truncating mutations in the *GRIN1*, *GRIN2A*, *GRIN2B* and *GRIN2D* genes in 854

individuals including 429 patients with schizophrenia, 142 with ASD and 283 healthy controls (285 controls from general population for *GRIN1*). By contrast, we identified the same truncating nonsense mutation in the *GRIN2C* gene (c.54G>A ; p.W18X) in two individuals. This mutation was first identified in an ASD patient and this variant was transmitted from his normal father while there was history of mild learning disabilities in the mother. We also found this mutation in a control from the general population

We also identified four truncating mutations in the *GRIN3A* gene. Two of them were indels creating frameshift in the open reading frame and creating nonsense codon; one in an individual from the population control cohort (c.2349delC ; p.P783PfsX23) and one in a healthy control (c.2424_2425delAA ; p.Q808QfsX5). The two other truncating mutations were single nucleotide variants. The first one (c.1522C>T ; p.Q508X) was identified in a patient diagnosed with catatonic schizophrenia and was inherited from a mother that had clinical characteristics ranging in the schizophrenia spectrum. The second one (c.679G>T ; p.E227X) was observed in a young control from the general population. Finally, we observed frequent (10%) truncating mutations in the *GRIN3B* gene in the same proportion as described elsewhere in the general population.⁵⁶

Overall, we identified 260 exonic variants in the seven sequenced NMDARs genes, including 159 variants that were not referenced in the dbSNP database (Build 132) (supplementary table 1 and 2). These variants were all confirmed manually by genotyping the proband and in the respective proband's parents, when available, to determine their inheritance. In order to avoid rare variants that could be linked to ethnic origin, we removed variants coming from patients with non-caucasian ethnicity. When doing so, we did not observe either accumulation of rare

variants in patients compared to controls or accumulation of more damaging variants in patients based on *in silico* analysis.

Discussion

Here, we provide the first systematic study of sequence alterations in all NMDARs genes in ASD and schizophrenia and report several potentially damaging and/or *de novo* mutations. Indeed, we identified potentially damaging *de novo* mutation in *GRIN2A* and *GRIN2B* associated with sporadic cases of schizophrenia and ASD, respectively, and truncating mutations in *GRIN2C*, *GRIN3A*, *GRIN3B*, whereas neither truncating nor *de novo* mutations were found in *GRIN1* and *GRIN2D*.

Deep sequencing technologies have brought new insight in the genetics of neurodevelopmental disorders including schizophrenia. Recent reports of the literature have shown that there should be, in average, less than two *de novo* mutations in the coding sequence per individual and less than one *de novo* that would have an effect on amino acid sequence per individual.^{44, 57-58}

We here report on three *de novo* mutations, two in *GRIN2A* in two patients suffering from schizophrenia and one in *GRIN2B*, in a patient with autism. The mutation in *GRIN2B* is one of the mutations quoted in a previous paper reporting an overall rate of *de novo* mutation in autism or schizophrenia.⁴⁴ In this former paper that considers the overall rate of *de novo* mutation, based on partial results, the mutation was unfortunately misclassified as being in a patient with schizophrenia while it is definitely in a patient with ASD. These mutations were not inherited from unaffected parents, and the patients with *de novo* mutation in *GRIN2A* or *GRIN2B* did not have a predisposing or causative mutation in 600 other synaptic genes that

have been sequenced, including schizophrenia risk genes, such as *DISC1* (data not shown). Hence, it could be hypothesized that these mutations could be associated with the pathogenesis of ASD or schizophrenia. This is further supported by the observation that the missense mutation in *GRIN2B* is predicted by *in silico* analysis to be relatively damaging and it occurred in the last trans-membrane domain that is highly conserved among species. In addition, one of the *de novo* mutations identified in the *GRIN2A* gene, is located in a domain that is highly conserved in mammals and vertebrates and is known to be the target of post-translational modifications, especially phosphorylation. In particular, it has been recently proposed that abnormalities in the neuregulin 1-ErbB4 pathway, that has been implicated in schizophrenia and known to contribute to NMDA receptor hypofunction, acts through NMDARs phosphorylation by the Src tyrosine kinase.¹⁸⁻¹⁹

Several studies have already screened for mutations in some NMDARs genes in neurodevelopmental diseases. Williams et al.⁵⁹ screened for mutations in the *GRIN1* and *GRIN2* genes in patients with schizophrenia but did not identify any variants that were clearly relevant to the phenotype with only a single missense variant identified among 368 individuals (including 184 controls). Endeley et al.²⁵ screened for mutations and cytogenetic abnormalities in the *GRIN2A* and *GRIN2B* genes in patients with ID and found rare damaging *de novo* mutations and truncating mutations in these genes indicating that these mutations can be pathogenic. Recently Hamdan et al., screened *GRIN1* in patients with non-syndromic ID and identified a *de novo* missense and one amino acid in-frame insertion that altered the receptor's activity,²⁶ the control population in this latter paper being the same as the one used in the present report. Lastly, an exome sequencing study published this year, identified several

potentially causative *de novo* mutations on patients with ASD, one of them being in *GRIN2B*, within a splice site.²⁹ Because we do not provide direct evidence of a functional effect of these mutations, we cannot definitely conclude on the causal effect of *de novo* mutations identified in *GRIN2A* and *GRIN2B* genes. It should however be emphasized that there were no truncating mutations in at least 854 individuals (including 283 healthy volunteers) in *GRIN1*, *GRIN2A*, *GRIN2B* and *GRIN2D*, supporting the hypothesis of a high selection pressure on these genes.

In addition to *de novo* mutations, we identified truncating mutations in the *GRIN2C*, *GRIN3A* and *GRIN3B* genes in patients, population controls and/or healthy volunteers. The rate of mutations in the *GRIN3B* gene was in the same proportion as described in the general population with around 10% of the individuals homozygous for truncating mutation⁵⁶ Because this relatively high rate of mutation and because *GRIN3B* is mainly expressed in motor neurons, and thus a role in mental disorders as schizophrenia and ASD is unlikely, the gene was not further explored in controls.^{56, 60}

The presence of truncating mutations in controls suggests that loss of function of these subunits leads to little or no effect. For the *GRIN2C* subunit, we identified the same truncating mutation in a patient with autism and in a control. Although the global functioning and learning abilities of this control appear to be in the normal range, his relatively young age at the time of his evaluation precludes from reaching firm conclusions on the effect of this truncating mutation on his future development.

On the other hand, the loss of function in one allele from the *GRIN3* family genes do not appear to have dramatic consequences as truncating mutations were identified in healthy

controls suggesting some functional compensations. NMDAR containing GRIN2 subunits show large differences in their pharmacological and electrophysiological properties, including glutamate affinity, modulation by glycine, sensitivity to Mg^{2+} and channel kinetics. Since they also show a distinct spatial and chronological pattern of expression, it has been suggested that these receptors may play unique roles in brain development.¹ Unlike most NMDAR that contain GRIN2 subunits, which typically require glycine and glutamate for activation, NMDAR that contain both GRIN1 and GRIN3 subunits are activated by glycine alone. When coexpressed with GRIN1 and GRIN2 subunits, GRIN3 subunits can act as negative modulators, reducing single-channel conductance and Ca^{2+} permeability.⁶¹ Thus, it is not surprising that rare damaging mutations could have differential effects in the GRIN2 and GRIN3 family of NMDARs.

We identified a large number of rare variants that are not in the dbSNP databases. Based on *in silico* analysis these variants do not appear to be pathogenic on their own. Further, most of them are inherited from unaffected parents and some of them could be linked to the ethnic origin since they were more frequent in the Tunisian samples. When excluding the mutations found in subjects that were not of European Caucasian origin, we did not observe any accumulation of rare or more damaging variants in patients compared to controls. Nevertheless, we cannot exclude complex gene interaction that could occur with variants in other genes.

In conclusion, our data support the hypothesis that rare *de novo* mutations in *GRIN2A* or *GRIN2B* could account for some cases of sporadic schizophrenia or autism. They also further support the implication of NMDA receptors in psychiatric disorders with neurodevelopmental

origin. Nevertheless, a distinct pattern is observed between the different NMDA receptor subunits. Truncating mutations in *GRIN2C* and *GRIN3* were found in patients and controls while no truncating mutations were found in *GRIN1*, *GRIN2A*, *GRIN2B* and *GRIN2D*. This could indicate that *GRIN1*, *GRIN2A*, *GRIN2B* and *GRIN2D* families are crucial in normal brain development and function and / or that functional compensation could occur to counteract the loss of one allele in *GRIN2C* and *GRIN3* family genes.

ACKNOWLEDGMENTS.

We would like to thank all the families involved in this study. This work was supported by Genome Canada and Génome Québec, and received co-funding from Université de Montréal for the ‘Synapse to Disease’ (S2D) project as well as funding from the Canadian Foundation for Innovation. J.T. receives a scholarship from the Université Paris Descartes (joint program Université Paris Descartes – Université de Montréal). G.A.R. holds the Canada Research Chair in Genetics of the Nervous System and a Jeanne-et-J.-Louis-Levesque Chair for the Genetics of Brain Diseases. P.D. holds the Canada Research Chair in Neuroscience. J.L.M. is a recipient of the Clinical Investigator Award of CIHR (Canadian Institutes of Health Research). The S2D team is composed of the following additional members: Isabelle Bachand, Marjolaine Chicoine, Mélanie Côté, Kathleen Daignault, Anne Desjarlais, Ousmane Diallo, Sylvia Dobrzeniecka, Joannie Duguay, Marina Drits, Philippe Jolivet, Liliane Karamera, Frédéric Kuku, Karine Lachapelle, Guy Laliberté, Sandra Laurent, Annie Levert, Meijiang Liao, Claude Marineau, Carlos Marino, Anne Noreau, Huashan Peng, Annie Raymond, Annie Reynolds, Daniel Rochefort, Judith St-Onge, Pascale Thibodeau, Kazuya Tsurudome, Yan Yang, Sophie Leroy, Narjes Bendjemaa, Katia Ossian, Mélanie Chayet, Fayçal Mouaffak. We would also like to thank Dr. Chawki Benkelfat for several helpful discussions. A portion of the Schizophrenia cohort was collected through the Collaborative Network for Family Study in Psychiatry ("Réseau d'étude familiale en Psychiatrie", REFAPSY), supported by the Fondation Pierre Deniker. We also acknowledge the efforts of the members of the Génome Québec Innovation Centre Sequencing (Pierre Lepage, Sébastien Brunet and Hao Fan Yam) and Bioinformatic (Louis Létourneau and Louis Dumond Joseph) groups.

Financial disclosures : The authors declare no conflict of interest

Bibliographie

1. Cull-Candy SG, Leszkiewicz DN. Role of distinct NMDA receptor subtypes at central synapses. *Sci STKE* 2004; **2004**(255): re16.
2. Krystal JH, Karper LP, Seibyl JP, Freeman GK, Delaney R, Bremner JD *et al.* Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry* 1994; **51**(3): 199-214.
3. Krystal JH, Anand A, Moghaddam B. Effects of NMDA receptor antagonists: implications for the pathophysiology of schizophrenia. *Arch Gen Psychiatry* 2002; **59**(7): 663-664.
4. Amitai N, Markou A. Disruption of performance in the five-choice serial reaction time task induced by administration of N-methyl-D-aspartate receptor antagonists: relevance to cognitive dysfunction in schizophrenia. *Biol Psychiatry* 2010; **68**(1): 5-16.
5. Vincent A, Bien CG. Anti-NMDA-receptor encephalitis: a cause of psychiatric, seizure, and movement disorders in young adults. *Lancet Neurol* 2008; **7**(12): 1074-1075.
6. Dalmau J, Gleichman AJ, Hughes EG, Rossi JE, Peng X, Lai M *et al.* Anti-NMDA-receptor encephalitis: case series and analysis of the effects of antibodies. *Lancet Neurol* 2008; **7**(12): 1091-1098.
7. Dalmau J, Lancaster E, Martinez-Hernandez E, Rosenfeld MR, Balice-Gordon R. Clinical experience and laboratory investigations in patients with anti-NMDAR encephalitis. *Lancet Neurol* 2011; **10**(1): 63-74.
8. Belforte JE, Zsiros V, Sklar ER, Jiang Z, Yu G, Li Y *et al.* Postnatal NMDA receptor ablation in corticolimbic interneurons confers schizophrenia-like phenotypes. *Nat Neurosci* 2010; **13**(1): 76-83.
9. Meador-Woodruff JH, Healy DJ. Glutamate receptor expression in schizophrenic brain. *Brain Res Brain Res Rev* 2000; **31**(2-3): 288-294.
10. Gao XM, Sakai K, Roberts RC, Conley RR, Dean B, Tamminga CA. Ionotropic glutamate receptors and expression of N-methyl-D-aspartate receptor subunits in

- subregions of human hippocampus: effects of schizophrenia. *Am J Psychiatry* 2000; **157**(7): 1141-1149.
11. Ibrahim HM, Hogg AJ, Jr., Healy DJ, Haroutunian V, Davis KL, Meador-Woodruff JH. Ionotropic glutamate receptor binding and subunit mRNA expression in thalamic nuclei in schizophrenia. *Am J Psychiatry* 2000; **157**(11): 1811-1823.
 12. Watis L, Chen SH, Chua HC, Chong SA, Sim K. Glutamatergic abnormalities of the thalamus in schizophrenia: a systematic review. *J Neural Transm* 2008; **115**(3): 493-511.
 13. Pilowsky LS, Bressan RA, Stone JM, Erlandsson K, Mulligan RS, Krystal JH *et al.* First in vivo evidence of an NMDA receptor deficit in medication-free schizophrenic patients. *Mol Psychiatry* 2006; **11**(2): 118-119.
 14. Kim JS, Kornhuber HH, Schmid-Burgk W, Holzmüller B. Low cerebrospinal fluid glutamate in schizophrenic patients and a new hypothesis on schizophrenia. *Neurosci Lett* 1980; **20**(3): 379-382.
 15. Akbarian S, Sucher NJ, Bradley D, Tafazzoli A, Trinh D, Hetrick WP *et al.* Selective alterations in gene expression for NMDA receptor subunits in prefrontal cortex of schizophrenics. *J Neurosci* 1996; **16**(1): 19-30.
 16. Allen NC, Bagade S, McQueen MB, Ioannidis JP, Kavvoura FK, Khoury MJ *et al.* Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* 2008; **40**(7): 827-834.
 17. Banerjee A, Macdonald ML, Borgmann-Winter KE, Hahn CG. Neuregulin 1-erbB4 pathway in schizophrenia: From genes to an interactome. *Brain Res Bull* 2010.
 18. Hahn CG, Wang HY, Cho DS, Talbot K, Gur RE, Berrettini WH *et al.* Altered neuregulin 1-erbB4 signaling contributes to NMDA receptor hypofunction in schizophrenia. *Nat Med* 2006; **12**(7): 824-828.
 19. Pitcher GM, Kalia LV, Ng D, Goodfellow NM, Yee KT, Lambe EK *et al.* Schizophrenia susceptibility pathway neuregulin 1-ErbB4 suppresses Src upregulation of NMDA receptors. *Nat Med* 2011.
 20. Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I *et al.* Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009; **460**(7256): 753-757.
 21. Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D *et al.* Common variants conferring risk of schizophrenia. *Nature* 2009; **460**(7256): 744-747.

22. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**(7256): 748-752.
23. Shatz CJ. MHC class I: an unexpected role in neuronal plasticity. *Neuron* 2009; **64**(1): 40-45.
24. Fourgeaud L, Davenport CM, Tyler CM, Cheng TT, Spencer MB, Boulanger LM. MHC class I modulates NMDA receptor function and AMPA receptor trafficking. *Proc Natl Acad Sci U S A* 2010; **107**(51): 22278-22283.
25. Endelev S, Rosenberger G, Geider K, Popp B, Tamer C, Stefanova I *et al.* Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat Genet* 2010; **42**(11): 1021-1026.
26. Hamdan FF, Gauthier J, Araki Y, Lin DT, Yoshizawa Y, Higashi K *et al.* Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. *Am J Hum Genet* 2011; **88**(3): 306-316.
27. Li R, Huang FS, Abbas AK, Wigstrom H. Role of NMDA receptor subtypes in different forms of NMDA-dependent synaptic plasticity. *BMC Neurosci* 2007; **8**: 55.
28. Gai X, Xie HM, Perin JC, Takahashi N, Murphy K, Wenocur AS *et al.* Rare structural variation of synapse and neurotransmission genes in autism. *Mol Psychiatry* 2011.
29. O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* 2011.
30. Eadie BD, Cushman J, Kannangara TS, Fanselow MS, Christie BR. NMDA receptor hypofunction in the dentate gyrus and impaired context discrimination in adult Fmr1 knockout mice. *Hippocampus* 2010.
31. Maliszewska-Cyna E, Bawa D, Eubanks JH. Diminished prevalence but preserved synaptic distribution of N-methyl-D-aspartate receptor subunits in the methyl CpG binding protein 2(MeCP2)-null mouse brain. *Neuroscience* 2010; **168**(3): 624-632.
32. Rinaldi T, Kulangara K, Antonello K, Markram H. Elevated NMDA receptor levels and enhanced postsynaptic long-term potentiation induced by prenatal exposure to valproic acid. *Proc Natl Acad Sci U S A* 2007; **104**(33): 13501-13506.
33. Blundell J, Blaiss CA, Etherton MR, Espinosa F, Tabuchi K, Walz C *et al.* Neuroligin-1 deletion results in impaired spatial memory and increased repetitive behavior. *J Neurosci* 2010; **30**(6): 2115-2129.

34. Polleux F, Lauder JM. Toward a developmental neurobiology of autism. *Ment Retard Dev Disabil Res Rev* 2004; **10**(4): 303-317.
35. Chez MG, Burton Q, Dowling T, Chang M, Khanna P, Kramer C. Memantine as adjunctive therapy in children diagnosed with autistic spectrum disorders: an observation of initial clinical response and maintenance tolerability. *J Child Neurol* 2007; **22**(5): 574-579.
36. Ng MY, Levinson DF, Faraone SV, Suarez BK, DeLisi LE, Arinami T *et al.* Meta-analysis of 32 genome-wide linkage studies of schizophrenia. *Mol Psychiatry* 2009; **14**(8): 774-785.
37. Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet* 2008; **40**(7): 880-885.
38. Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R *et al.* Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 2010; **466**(7304): 368-372.
39. Guilmatre A, Dubourg C, Mosca AL, Legallic S, Goldenberg A, Drouin-Garraud V *et al.* Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. *Arch Gen Psychiatry* 2009; **66**(9): 947-956.
40. Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P *et al.* A de novo paradigm for mental retardation. *Nat Genet* 2010; **42**(12): 1109-1112.
41. Hamdan FF, Daoud H, Rochefort D, Piton A, Gauthier J, Langlois M *et al.* De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. *Am J Hum Genet* 2010; **87**(5): 671-678.
42. Hamdan FF, Gauthier J, Spiegelman D, Noreau A, Yang Y, Pellerin S *et al.* Mutations in SYNGAP1 in autosomal nonsyndromic mental retardation. *N Engl J Med* 2009; **360**(6): 599-605.
43. Gauthier J, Siddiqui TJ, Huashan P, Yokomaku D, Hamdan FF, Champagne N *et al.* Truncating mutations in NRXN2 and NRXN1 in autism spectrum disorders and schizophrenia. *Hum Genet* 2011.
44. Awadalla P, Gauthier J, Myers RA, Casals F, Hamdan FF, Griffing AR *et al.* Direct Measure of the De Novo Mutation Rate in Autism and Schizophrenia Cohorts. *Am J Hum Genet* 2010.

45. Tarabeux J. De Novo Truncating Mutation in Kinesin 17 Associated with Schizophrenia. *Biol Psychiatry* 2010.
46. Gauthier J, Champagne N, Lafreniere RG, Xiong L, Spiegelman D, Brustein E *et al.* De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for schizophrenia. *Proc Natl Acad Sci U S A* 2010; **107**(17): 7863-7868.
47. Knight HM, Pickard BS, Maclean A, Malloy MP, Soares DC, McRae AF *et al.* A cytogenetic abnormality and rare coding variants identify ABCA13 as a candidate gene in schizophrenia, bipolar disorder, and depression. *Am J Hum Genet* 2009; **85**(6): 833-846.
48. Setou M, Nakagawa T, Seog DH, Hirokawa N. Kinesin superfamily motor protein KIF17 and mLin-10 in NMDA receptor-containing vesicle transport. *Science* 2000; **288**(5472): 1796-1802.
49. Rapoport JL, Addington A, Frangou S. The neurodevelopmental model of schizophrenia: what can very early onset cases tell us? *Curr Psychiatry Rep* 2005; **7**(2): 81-82.
50. Gourion D, Goldberger C, Olie JP, Loo H, Krebs MO. Neurological and morphological anomalies and the genetic liability to schizophrenia: a composite phenotype. *Schizophr Res* 2004; **67**(1): 23-31.
51. DeLisi LE, Shaw SH, Crow TJ, Shields G, Smith AB, Larach VW *et al.* A genome-wide scan for linkage to chromosomal regions in 382 sibling pairs with schizophrenia or schizoaffective disorder. *Am J Psychiatry* 2002; **159**(5): 803-812.
52. Joober R, Rouleau GA, Lal S, Dixon M, O'Driscoll G, Palmour R *et al.* Neuropsychological impairments in neuroleptic-responder vs. -nonresponder schizophrenic patients and healthy volunteers. *Schizophr Res* 2002; **53**(3): 229-238.
53. Fathalli F, Rouleau GA, Xiong L, Tabbane K, Benkelfat C, Deguzman R *et al.* No association between the DRD3 Ser9Gly polymorphism and schizophrenia. *Schizophr Res* 2008; **98**(1-3): 98-104.
54. Piton A, Michaud JL, Peng H, Aradhya S, Gauthier J, Motttron L *et al.* Mutations in the calcium-related gene IL1RAPL1 are associated with autism. *Hum Mol Genet* 2008; **17**(24): 3965-3974.
55. Lemelin JP, Boivin M, Forget-Dubois N, Dionne G, Seguin JR, Brendgen M *et al.* The genetic-environmental etiology of cognitive school readiness and later academic achievement in early childhood. *Child Dev* 2007; **78**(6): 1855-1869.

56. Niemann S, Landers JE, Churchill MJ, Hosler B, Sapp P, Speed WC *et al.* Motoneuron-specific NR3B gene: no association with ALS and evidence for a common null allele. *Neurology* 2008; **70**(9): 666-676.
57. Nachman MW, Crowell SL. Estimate of the mutation rate per nucleotide in humans. *Genetics* 2000; **156**(1): 297-304.
58. Lynch M. Rate, molecular spectrum, and consequences of human mutation. *Proc Natl Acad Sci U S A* 2010; **107**(3): 961-968.
59. Williams NM, Bowen T, Spurlock G, Norton N, Williams HJ, Hoogendoorn B *et al.* Determination of the genomic structure and mutation screening in schizophrenic individuals for five subunits of the N-methyl-D-aspartate glutamate receptor. *Mol Psychiatry* 2002; **7**(5): 508-514.
60. Niemann S, Kanki H, Fukui Y, Takao K, Fukaya M, Hynynen MN *et al.* Genetic ablation of NMDA receptor subunit NR3B in mouse reveals motoneuronal and nonmotoneuronal phenotypes. *Eur J Neurosci* 2007; **26**(6): 1407-1420.
61. Das S, Sasaki YF, Rothe T, Premkumar LS, Takasu M, Crandall JE *et al.* Increased NMDA current and spine density in mice lacking the NMDA receptor subunit NR3A. *Nature* 1998; **393**(6683): 377-381.

Table 1 : Genes and samples screened

	samples			
	SCZ	ASD	Controls QNTS	Controls
<i>GRIN1</i>	429 subjects (292 trios)	142 trios	285 trios	<i>not analyzed</i>
<i>GRIN2A</i>	429 subjects (292 trios)	142 trios	<i>not analyzed</i>	283 subjects
<i>GRIN2B</i>	429 subjects (292 trios)	428 subjects (404 trios)	<i>not analyzed</i>	283 subjects
<i>GRIN2C</i>	429 subjects (292 trios)	428 subjects (404 trios)	285 trios	283 subjects
<i>GRIN2D</i>	429 subjects (292 trios)	142 trios	<i>not analysed</i>	283 subjects
<i>GRIN3A</i>	429 subjects (292 trios)	142 trios	285 trios	283 subjects
<i>GRIN3B</i>	429 subjects (292 trios)	428 subjects (404 trios)	<i>not analyzed</i>	<i>not analyzed</i>

Abbreviations : SCZ, schizophrenia; ASD, autism spectrum disorders; QNTS, Quebec newborn twin study.

Number of trios indicates individuals having parental DNA from both parents available

Table 2 : Details of de novo mutations identified in NMDAR subunits

Gene	sequence change						samples		<i>In silico</i> prediction for amino acid	
	Variant type	Protein change	Chr	Chr Position ¹	Exon	cDNA change	SCZ	ASD	SIFT	Polyphen
GRIN2A	Missense	p.A968T	16	9858499	14	c.2902G>A	1 (292 trios)	0 (142 trios)	0,15 (tolerated)	0,069 (benign)
	Silent	p.T1223T	16	9857732	14	c.3669C>T	1 (292 trios)	0 (142 trios)	/	/
GRIN2B	Missense	p.L825V	12	13720084	12	c.2473T>G	0 (292 trios)	1 (404 trios)	0,18 (tolerated)	1,567 (possibly damaging)

Abbreviations : chr, chromosome; SCZ, schizophrenia; ASD, autism spectrum disorders

¹Positions according to genome build 37

Number of trios indicates individuals having parental DNA from both parents available

PolyPhen <http://genetics.bwh.harvard.edu/pph>

SIFT <http://blocks.fhcrc.org/sift/SIFT.html>

SNAP <http://www.rostlab.org/services/SNAP>

Table 3 : Details of rare truncating mutations in NMDAR subunits

Gene	sequence change						samples		
	Variant type	protein change	Chr	Chr Position ¹	exon	cDNA change	ASD frequency	SCZ frequency	controls frequency
GRIN2C	Nonsense	p.W18X	17	72851178	1	c.54G>A	1/428	0/429	1/568
GRIN3A	Nonsense	p.E227X	9	104499583	1	c.679G>T	0/142	0/429	1/568
	Nonsense	p.Q508X	9	104433172	3	c.1522C>T	0/142	1/429	0/568
	Indel	p.P783PfsX23	9	104432345	3	c.2349delC	0/142	0/429	1/568
	Indel	p.Q808QfsX6	9	104390612_104390613	4	c.2424_2425delAA	0/142	0/429	1/568

Abbreviations : chr, chromosome; SCZ, schizophrenia; ASD, autism spectrum disorders

¹Positions according to genome build 37

Supplementary table 1: *GRIN1* and *GRIN2* exonic nucleotide changes not annotated in dbSNP.

Gene	sequence alteration					samples				In silico prediction for amino acid substitution			
	Variant Type	Protein Change	Chr. Position	cDNA position	SCZ (n=429)	ASD (n=142)	Controls QNTS (n=285)	Controls (n=283)	ethnicity	Pathfinder	Sift	PolyPhen	
GRIN1	MISSENSE	p.A309Q	140051438	c.917C>A	1	0	0	0	caucasian	-3.86931	0.51	1.353	
	MISSENSE	p.A349S	140052907	c.1045G>T	1	0	0	0	caucasian	-2.85961	0.42	1.347	
	MISSENSE	p.T419A	140055565	c.1255A>G	0	1	0	0	caucasian	-3.24749	0.94	1.200	
	SILENT	p.V231V	140051142	c.693A>C	0	0	1	0	caucasian	/	/	/	
GRIN2A	MISSENSE	p.T189N	10032257	c.566C>A	1	0	0	0	caucasian	-4.41615	0.16	1.379	
	MISSENSE	p.V452M	9934936	c.1354G>A	2	0	0	0	African	-5.14524	0.07	0.537	
	MISSENSE	p.V852M	9862749	c.2554G>A	1	0	0	0	African	NA	0.04	0.099	
	MISSENSE	p.V998M	9858409	c.2992G>A	1	0	0	0	caucasian	NA	0.04	0.977	
	MISSENSE	p.T1064A	9858211	c.3190A>G	1	0	0	0	caucasian	NA	0.63	0.180	
	MISSENSE	p.T1229S	9857716	c.3685A>T	1	0	0	0	caucasian	NA	0.38	0.422	
	MISSENSE	p.R1159P	9857925	c.3788G>C	1	0	0	0	Tunisian	NA	0.01	1.774	
	SILENT	p.P273P	10032004	c.819A>G	1	0	0	0	caucasian	/	/	/	
	SILENT	p.V417V	9943690	c.1561C>T	1	0	0	0	Tunisian	/	/	/	
	SILENT	p.V734V	9892288	c.2512C>G	1	0	0	0	African	/	/	/	
	SILENT	p.T880T	9858737	c.2974G>A	1	0	0	0	African	/	/	/	
	SILENT	p.V1139V	9857984	c.3727G>T	1	0	0	0	caucasian	/	/	/	
	SILENT	p.P1142P	9857975	c.3736C>T	1	0	0	0	Tunisian	/	/	/	
	MISSENSE	p.T143I	10032395	c.428C>T	0	1	0	0	caucasian	-4.72691	0.38	0.412	
	MISSENSE	p.A922V	9858636	c.2765C>T	0	1	0	0	caucasian	NA	0.42	0.751	
	SILENT	p.L63L	10274082	c.501C>T	0	1	0	0	caucasian	/	/	/	
	MISSENSE	p.P336S	10031817	c.1006C>T	0	0	0	1	caucasian	-3.67251	0.49	1.085	
	MISSENSE	p.F576S	9928012	c.1727T>C	0	0	0	1	Tunisian	-4.7378	0.02	2.576	
	MISSENSE	p.D937N	9858592	c.2809G>A	0	0	0	1	caucasian	NA	0.43	1.130	
	MISSENSE	p.A1276G	9857574	c.3827C>G	0	0	0	2	caucasian	NA	0.37	0.335	
	SILENT	p.D234D	10032121	c.702C>T	0	0	0	1	caucasian	/	/	/	
	SILENT	p.T470T	9934880	c.1410T>G	0	0	0	1	Tunisian	/	/	/	
	MISSENSE	p.V18I	14019001	c.52G>A	2	0*	0	0	caucasian	-3.50328	1.00	0.673	
	MISSENSE	p.I50N	14018994	c.149T>A	1	0*	0	0	caucasian	-5.56556	0.03	1.526	
	MISSENSE	p.L362M	13828720	c.1084C>A	1	0*	0	0	caucasian	-3.48707	0.08	0.380	
	MISSENSE	p.Q1014R	13717131	c.3041A>G	1	0*	0	0	caucasian	NA	0.83	0.339	
GRIN2B	INDEL(in frame)	p.N1352del	13716114, 13716116	c.4056_4058delCAA	1	0*	0	0	caucasian	/	/	/	
	MISSENSE	p.L1424F	13715902	c.4270C>T	1	0*	0	0	caucasian	NA	0.34	0.394	
	SILENT	p.S155S	13906796	c.465C>A	1	0*	0	0	African	/	/	/	
	SILENT	p.Y452Y	13768571	c.1356C>T	1	0*	0	0	caucasian	/	/	/	
	SILENT	p.S617S	13761696	c.1851C>T	2	0*	0	0	caucasian	/	/	/	
	SILENT	p.A734A	13722921	c.2202A>G	1	0*	0	0	African	/	/	/	
	SILENT	p.I1105I	13716857	c.3315C>T	1	0*	0	0	caucasian	/	/	/	
	SILENT	p.R1164R	13716680	c.3492C>T	1	0*	0	0	caucasian	/	/	/	
	SILENT	p.P1219P	13716515	c.3657C>T	1	0*	0	0	caucasian	/	/	/	
	MISSENSE	p.G1026S	13717096	c.3076G>A	0	1*	0	0	caucasian	NA	0.92	0.044	
	MISSENSE	p.S1415L	13715928	c.4244C>T	0	1*	0	0	caucasian	NA	0.01	1.533	
	SILENT	p.V1308V	13716248	c.3924G>A	0	1*	0	0	caucasian	/	/	/	
	MISSENSE	p.M1342R	13716147	c.4025T>G	0	0*	0	1	caucasian	NA	0.07	0.881	
	SILENT	p.T685T	13724854	c.2055C>T	0	0*	0	1	Tunisian	/	/	/	
	SILENT	p.Y1371Y	13716059	c.4113C>T	0	0*	0	1	caucasian	/	/	/	
	SILENT	p.G1471G	13715759	c.4413G>A	0	0*	0	1	caucasian	/	/	/	
	MISSENSE	p.A271V	13906449	c.812C>T	1	0*	0	1	caucasian	-2.95627	0.39	1.113	
	MISSENSE	p.S1452F	13715817	c.4355C>T	1	0*	0	1	caucasian	NA	0.10	1.222	
	SILENT	p.G1184G	13716620	c.3552C>T	1	3*	0	0	caucasian	/	/	/	
GRIN2C	MISSENSE	p.A166V	72848653	c.497C>T	1	0*	0	0	caucasian	-3.36354	0.55	0.01	
	MISSENSE	p.A641T	72843527	c.1821G>A	1	0*	0	0	caucasian	-5.38822	0.00	1.542	
	MISSENSE	p.R679C	72843026	c.2035C>T	1	0*	0	0	filipino	-7.10378	0.01	2.564	
	MISSENSE	p.S911I	72839544	c.2732G>T	1	0*	0	0	affrican	NA	0.05	1.875	
	MISSENSE	p.S989P	72839311	c.2965T>C	1	0*	0	0	caucasian	NA	1.00	0.442	
	MISSENSE	p.T1196I	72838689	c.3587C>T	1	0*	0	0	Tunisian	NA	0.08	1.35	
	SILENT	p.L215L	72848507	c.643C>T	1	0*	0	0	caucasian	/	/	/	
	SILENT	p.Y473Y	72846417	c.1419C>T	1	0*	0	0	caucasian	/	/	/	
	SILENT	p.P1024P	72839204	c.3072C>T	1	0*	0	0	caucasian	/	/	/	
	INDEL (in frame)	p.G244_L245insP	72848418, 72848419	c.733_734insCCG	0	1*	0	0	caucasian	/	/	/	
	MISSENSE	p.I863T	72839688	c.2588T>C	0	1*	0	0	caucasian	NA	0.12	1.824	
	SILENT	p.T131T	72850839	c.393C>T	0	1*	0	0	caucasian	/	/	/	
	SILENT	p.A264A	72846358	c.792G>A	0	1*	0	0	caucasian	/	/	/	
	SILENT	p.V759V	72842278	c.2277C>T	0	1*	0	0	caucasian	/	/	/	
	SILENT	p.S1018S	72839222	c.3054C>A	0	2*	0	0	caucasian	/	/	/	
	MISSENSE	p.R537C	72845955	c.1609C>T	0	0*	0	1	caucasian	-6.18782	0.01	2.654	
	SILENT	p.A238A	72848436	c.714C>G	0	0*	1	0	caucasian	/	/	/	
	SILENT	p.N803N	72840589	c.2409C>T	0	0*	1	0	caucasian	/	/	/	
	SILENT	p.P882P	72839630	c.2646G>A	0	0*	1	0	caucasian	/	/	/	
	MISSENSE	p.V90I	72850964	c.268G>A	0	0*	0	1	caucasian	-2.87339	0.17	0.159	
	MISSENSE	p.M573T	72844044	c.1718T>C	0	0*	0	1	caucasian	-3.26651	0.02	1.933	
	MISSENSE	p.Q871R	72839664	c.2612A>G	0	0*	0	1	caucasian	NA	0.21	1.434	
	MISSENSE	p.P982A	72839332	c.2944C>G	0	0*	0	1	caucasian	NA	1.00	0.557	
	MISSENSE	p.H1079Y	72839041	c.3235C>T	0	0*	0	1	caucasian	NA	0.01	1.350	
	MISSENSE	p.P1141A	72838855	c.3421C>G	0	0*	0	1	Tunisian	NA	0.94	1.125	
	SILENT	p.T570T	72844052	c.1947C>T	0	0*	0	1	caucasian	/	/	/	
	SILENT	p.S710S	72842931	c.2130G>A	1	0*	0	1	caucasian	/	/	/	
	SILENT	p.P955P	72839411	c.2885G>A	0	0*	0	2	caucasian	/	/	/	
	MISSENSE	p.P877S	72839647	c.2629C>T	0	1*	0	2	caucasian	NA	0.75	0.206	
	MISSENSE	p.S995L	72839292	c.2984C>T	1	1*	0	0	caucasian	NA	0.01	1.350	
	MISSENSE	p.G1180R	72838738	c.3538G>C	2	1*	0	4	caucasian	NA	0.05	1.800	
GRIN2D	INDEL (in frame)	p.H1187_G1194delinst.	72838696, 72838716	c.3560_3580del	1	0*	0	2	caucasian	/	/	/	
	SILENT	p.P222P	72848484	c.666C>T	1	1*	0	0	caucasian	/	/	/	
	SILENT	p.N480N	72846396	c.1440C>T	0	1*	0	1	caucasian	/	/	/	
	MISSENSE	p.R237L	48908235	c.653G>T	1	0	0	0	pakisthanese	-5.3682	0.28	1.903	
	MISSENSE	p.L466V	48917825	c.1396C>G	1	0	0	0	tunisian	-2.91396	0.49	0.114	
	MISSENSE	p.M733V	48925147	c.2197A>G	1	0	0	0	caucasian	-5.06581	0.06	2.266	
	MISSENSE	p.R872H	48945581	c.2615G>A	1	0	0	0	caucasian	NA	0.12	0.209	
	MISSENSE	p.M883I	48945615	c.2649G>A	1	0	0	0	caucasian	NA	0.29	1.509	
	MISSENSE	p.A901I	48945884, 48945885	c.2701G>A+2702C>T	1	0	0	0	tunisian	NA	0.12	0.900	
	MISSENSE	p.A922V	48945948	c.2765C>T	2	0	0	0	African	NA	0.6	0.900	
	MISSENSE	p.A982P	48946127	c.2944G>C	1	0	0	0	asian	NA	0.07	1.125	
	SILENT	p.P225P	48908200	c.675T>G	1	0	0	0	African	/	/	/	
	SILENT	p.A1284A	48947035	c.3852G>C	1	0	0	0	caucasian	/	/	/	
	MISSENSE	p.M592L	48922529	c.1774A>C	0	1	0	0	caucasian	-3.481	0.51	2.041	
	MISSENSE	p.A926T	48945959	c.2776G>A	0	1	0	0	caucasian	NA	0.11	0.000	
	SILENT	p.N522N	48918274	c.1566C>T	0	1	0	0	caucasian	/	/	/	
	SILENT	p.T393T	4891734	c.1179C>T	0	0	0	1	tunisian	/	/	/	
	MISSENSE	p.G1317S	48947132	c.3949G>A	1	2	0	0	mixed	NA	0.02	NA	
	SILENT	p.I757I	48945044	c.2271C>T	1	3	0	0	caucasian	/	/	/	

* n=428 patients with ASD

Supplementary table 2 : *GRIN3* exonic nucleotide changes not annotated in dbSNP.

Gene	sequence alteration				samples				<i>In silico</i> prediction for amino acid substitution			
	Variant Type	Protein Change	Chr. Position	cDNA position	SCZ (n=429)	ASD (n=142)	Controls QNTS (n=285)	Controls (n=283)	ethnicity	Panther	Sift	PolyPhen
GRIN3A	MISSENSE	p.S72I	104500047	c.215G>T	2	0	0	0	caucasian	NA	0.05	1.592
	MISSENSE	p.Y493F	104433216	c.1478A>T	1	0	0	0	caucasian	-3.49689	0.09	1.626
	MISSENSE	p.I1028N	104335721	c.3083T>A	1	0	0	0	caucasian	NA	0.52	1.175
	SILENT	p.A119A	104499905	c.357G>A	1	0	0	0	caucasian	/	/	/
	SILENT	p.F199F	104499665	c.597C>T	1	0	0	0	Iranian	/	/	/
	SILENT	p.C330C	104449192	c.990C>T	1	0	0	0	caucasian	/	/	/
	SILENT	p.S888S	104375760	c.2664C>T	1	0	0	0	caucasian	/	/	/
	SILENT	p.T921T	104375661	c.2763G>A	1	0	0	0	Indian	/	/	/
	SILENT	p.L979L	104340121	c.2937A>G	2	0	0	0	caucasian	/	/	/
	SILENT	p.V1000V	104340058	c.3000G>A	1	0	0	0	afican	/	/	/
	MISSENSE	p.R111G	104499931	c.331C>G	0	1	0	0	caucasian	NA	0.13	0.474
	MISSENSE	p.G373R	104449065	c.1117G>A	0	3	0	0	caucasian	-4.85999	0.06	2.274
	MISSENSE	p.V641L	104432773	c.1921G>T	0	1	0	0	caucasian	-4.93781	0.06	1.430
	MISSENSE	p.R1009C	104435779	c.3026G>A	0	1	0	0	caucasian	NA	0.06	1.568
	MISSENSE	p.I375T	104449058	c.1124T>C	0	0	1	0	caucasian	-3.54425	0.01	1.618
	MISSENSE	p.S434P	104448882	c.1300T>C	0	0	1	0	caucasian	-2.93569	0.22	1.626
	SILENT	p.F572F	104432978	c.1716C>T	0	0	1	0	caucasian	/	/	/
	SILENT	p.Y805Y	104390621	c.2415T>C	0	0	1	0	caucasian	/	/	/
	MISSENSE	p.R337W	104449173	c.1009C>T	0	0	0	1	Tunisian	-6.93407	0.02	2.131
	MISSENSE	p.S586N	104432936	c.1758G>C	0	0	0	1	caucasian	-4.57512	0.09	1.541
	SILENT	p.T505T	104433179	c.1515C>T	0	0	0	1	Tunisian	/	/	/
	SILENT	p.D596D	104432906	c.1788C>T	0	0	0	1	caucasian	/	/	/
	SILENT	p.T761T	104432411	c.2283G>A	0	0	0	1	caucasian	/	/	/
	MISSENSE	p.T921M	104375662	c.2762C>T	0	0	1	2	caucasian	NA	0.17	2.050
	MISSENSE	p.G116S	104499916	c.346G>A	4	1	2	0	caucasian	-3.99333	0.47	0.810
	MISSENSE	p.G487V	104433234	c.1457G>T	1	1	0	1	caucasian	-2.93445	0.39	2.499
	MISSENSE	p.N565S	104433000	c.1694A>G	1	1	0	0	caucasian	-2.80401	0.03	0.894
	MISSENSE	p.R793S	104390659	c.2377C>A	1	0	0	1	caucasian	-4.21391	0.16	2.052
	MISSENSE	p.Y956H	104341543	c.2866T>C	1	0	0	1	caucasian	NA	0.33	0.110
	SILENT	p.F931F	104341616	c.2793C>T	1	0	2	0	caucasian	/	/	/
GRIN3B	MISSENSE	p.R194V	1003282	c.580C>T	1	0*	/	/	afican	-5.10873	0.07	2.217
	MISSENSE	p.R247Q	1003442	c.740G>A	1	0*	/	/	asian	-3.43248	0.61	0.124
	MISSENSE	p.R336C	1003708	c.1006C>T	1	0*	/	/	caucasian	-4.84314	0.03	1.849
	MISSENSE	p.R350H	1004549	c.1049G>A	1	0*	/	/	caucasian	-3.44963	0.28	1.698
	MISSENSE	p.P381L	1004642	c.1142C>T	1	0*	/	/	caucasian	-3.04547	0.79	2.846
	MISSENSE	p.R470H	1004909	c.1409G>A	1	0*	/	/	caucasian	-3.19818	0.42	1.532
	MISSENSE	p.A583V	1005248	c.1748C>T	1	0*	/	/	caucasian	-2.84435	1.00	0.317
	MISSENSE	p.R598C	1005292	c.1792C>T	2	0*	/	/	afican	-6.81203	0.18	2.097
	MISSENSE	p.R608H	1005323	c.1823G>A	1	0*	/	/	caucasian	NA	0.06	1.979
	MISSENSE	p.V928M	1009251	c.2782G>A	2	0*	/	/	afican	NA	0.11	1.004
	MISSENSE	p.A948V	1009312	c.2843C>T	1	0*	/	/	caucasian	NA	0.41	0.331
	NONSENSE	p.W575X	1005224	c.1724G>A	0	2*	/	/	caucasian	/	/	/
	MISSENSE	p.R180C	1003240	c.538C>T	0	2*	/	/	caucasian	-5.8923	0.08	0.399
	MISSENSE	p.R247W	1003441	c.739C>T	0	1*	/	/	caucasian	-6.93407	0.01	1.853
	MISSENSE	p.A285G	1003556	c.854C>G	0	1*	/	/	caucasian	-2.81006	0.41	0.389
	MISSENSE	p.E332K	1003696	c.994G>A	0	1*	/	/	caucasian	-3.86633	0.30	0.248
	MISSENSE	p.L472P	1004915	c.1415T>C	0	1*	/	/	caucasian	-4.16404	0.20	1.572
	MISSENSE	p.G515S	1005043	c.1543G>A	0	1*	/	/	caucasian	-6.72034	0.02	1.824
	MISSENSE	p.T641M	1005422	c.1922C>T	0	1*	/	/	caucasian	-4.64514	0.07	2.050
	MISSENSE	p.S678L	1005533	c.2033C>T	0	1*	/	/	caucasian	-5.16326	0.03	2.018
	MISSENSE	p.E919D	1009226	c.2757G>C	0	2*	/	/	caucasian	NA	0.63	0.113
	SILENT	p.A589A	1005267	c.1767G>A	0	1*	/	/	caucasian	/	/	/

* n=428 patients with ASD

Supplementary table 3 : Primers used for screening of the seven NMDA receptor subunit genes. Start and stop position are based on the human genome hg19 assembly

Gene	Primer name	forward	reverse	amplicon size	target refseq	exon	chromosome	start	stop
GRIN1	G00166_002	TGCTTCCAGATCTCAAGCCTC	CCCTGCTCTCTGGAGTG	286	AB208918	2	9	140036388	140036673
GRIN1	G00166_003	GTGTCTCGCAGTGGGAG	ACGGAGCCACAGACACAG	333	AB208918	3	9	140040098	140040430
GRIN1	G00166_004	TTAGCGCGCTATTTCAC	GCTGGGCCAGACAGTGG	225	AF015731	4	9	140042529	140042753
GRIN1	G00166_005	CAACTCTCAACCCCTGAGGC	CCCAACAATGCTCCAGCTC	260	AB208918	4	9	140043392	140043651
GRIN1	G00166_012	ACTGATTCGCCACCCAGGC	AGTATACAGCATGAGACCTC	360	AB208918	17	9	140057871	140058230
GRIN1	G00166_022	CCCTGCGACCTTACTTTTG	GACTCTGAGACAGAGGGGC	539	AB208918	18	9	140058089	140058627
GRIN1	G00166_014	ATGCAGGAGCAGGCGAG	GATGGGGCGGAGACAC	254	NM_007327	19	9	140059567	140059820
GRIN1	G00166_028	CTTTGGGTAGGGTGGGGTC	GTGCTCCGGGGCTCTG	360	NM_007327	20	9	140061713	140062072
GRIN1	G00166_024	GGAGAGCTGAGACTCCCG	CTTCCTCTAGGCTCTCC	407	AF015731	21	9	140061970	140062376
GRIN1	G00166_032	CGACCCCAAGATCTGCAAC	CCAGCTGACTCTGGGGTG	319	NM_007327	1,2	9	140034004	140034322
GRIN1	G00166_019	GAGCTGGGTAGGGCTTTGG	CTCTCTGCTGTCAACCTTG	637	AB208918	11-12	9	140056253	140056889
GRIN1	G00166_020	GATGCTCTAGCTCTGGAG	AACTCTGCTCGAGGGTCTG	606	AB208918	13-14	9	140056716	140057321
GRIN1	G00166_021	GAATCGAGGTGAGCTGG	GGAAGTGTCTGACATCTG	644	AB208918	15-16	9	140057218	140057861
GRIN1	G00166_018	GAGGAGGAGGACTGGGC	GCCCATCTCTGAGATGACC	520	AB208918	5-6	9	140051042	140051561
GRIN1	G00166_007	GGGGAAGGAGCAGGAG	TGGCTTTTATTGACCTCAC	495	AB208918	7-8	9	140052762	140053256
GRIN1	G00166_008	TGAGAAGAGACTGCCGCC	TACTCCATGCCGCCACCC	507	AB208918	9-10	9	140055439	140055945
GRIN2A	G00162_001	TTTCTTCTGCTGCCCTTG	CAAGACAGATTCTAGGGGGC	582	NM_000833	3	16	10274341	10273760
GRIN2A	G00162_002	TTCTAGGACGAGTTTGTGC	AGAATATACCCGGTGAAG	422	NM_000833	4	16	10032484	10032063
GRIN2A	G00162_003	AGATCCACTCTCTGCTCTTG	CAGTCGATTTTCCATATTTC	441	NM_000833	4	16	10032161	10031721
GRIN2A	G00162_004	TGTTGGCATCCACTATTTG	GAAGAAGCACTGTGAGCC	267	NM_000833	5	16	9985032	9984766
GRIN2A	G00162_005	GAAAGGATTTGCTCTCCAG	ATGACATTTTGTGATTCAGTAC	376	NM_000833	6	16	9943888	9943513
GRIN2A	G00162_007	CAATGGATCTCTGGGCTTC	TCAGTGACCCTCAAACTCTTC	640	NM_000833	9	16	9928169	9927530
GRIN2A	G00162_008	TGCTGCTGGTTGTCTATAC	AGAGGCACCTGAATCTCTTC	386	NM_000833	10	16	9923583	9923198
GRIN2A	G00162_009	TGCTATTTCTCTCAACACAG	GCATGCGAGAGTCAATTCG	353	NM_000833	11	16	9916400	9916048
GRIN2A	G00162_010	AGCTGGGATGCTTTCAGG	GGAGGCCACAGAGCAAC	368	NM_000833	12	16	9892415	9892059
GRIN2A	G00162_011	ATGCTGTGTCTCAAAACAG	AGACATCAAGAACCAAGCG	393	NM_000833	13	16	9852029	9851673
GRIN2A	G00162_012	GGTAAAGTCTAAGTCAAAATGG	TCTGGGAGTCTGATTTCTGG	370	NM_000833	14	16	9858881	9858301
GRIN2A	G00162_013	CGTGAGCAGAGATCCAAAG	ACTGGCTGTTGTGGAAAGC	554	NM_000833	14	16	9858404	9857851
GRIN2A	G00162_014	GACTCTCCGAGCCCTAC	TGCTTGGAGGGACACTAAAC	569	NM_000833	14	16	9857965	9857397
GRIN2A	G00162_015	TGTGACCAACCTAGGAGGC	TTGGAACCTCAACATTTACC	590	NM_000833	14	16	9857516	9856927
GRIN2A	G00162_006	GCTCTGGGGAAGCCACTTC	TTCTTTCTTGGTCCCTACG	675	NM_000833	7-8	16	9935093	9934419
GRIN2B	G00167_001	CGGTAAATGCGAGAGAAGC	TCCCAAAACCACTAATTTC	324	NM_000834	1	12	14133188	14132728
GRIN2B	G00167_001	GGAGTAAGGCTCTTGAATTC	GCTGTCAATGCAATCTGGTAC	644	NM_000834	2	12	14019256	14018613
GRIN2B	G00167_002	TGTTGTCAAATCTCAATACC	ATCCAGGTGATGACCTGACC	435	NM_000834	3	12	13906925	13906491
GRIN2B	G00167_003	TCAAGAACTTCAAGGCCCC	TTCTGAATCTCAACCCCAAG	428	NM_000834	3	12	13906602	13906175
GRIN2B	G00167_004	TCCTATATACCTGACTGCTTGTTC	CAAAAGCCAAAGGACTCTTTTC	283	NM_000834	4	12	13828770	13828588
GRIN2B	G00167_005	TCAACTGCTTCTCGGGATTG	AGCCAGTGTCTAACAGGTGCG	370	NM_000834	5	12	13769687	13769318
GRIN2B	G00167_006	AAATGCTGCTGAGCTGTGAAG	CAGCTCTTACAGTGGTTCTTTT	459	NM_000834	6	12	13768727	13768269
GRIN2B	G00167_007	CAAGAAAGCTGTATAGGCTG	TGAGCTCTTACGCAAAACAC	362	NM_000834	7	12	13768291	13767930
GRIN2B	G00167_008	ATTGTGGTTTGTCTCCACAC	CAATGAGGAGTCCAGAGATTG	308	NM_000834	8	12	13764892	13764585
GRIN2B	G00167_009	GGTTTATGCGGAAGGGTTTC	AGGGAACCTGACTCTCTCTCC	503	NM_000834	9	12	13761881	13761379
GRIN2B	G00167_010	GGGTCCTCTCAATATAACCTG	GAAATCCATAAAGACAATGAAG	429	NM_000834	10	12	13725026	13724598
GRIN2B	G00167_011	CATATTAATGGCCCAACAGC	GCACAGGTAAAGAAATGGAAG	356	NM_000834	11	12	13723024	13722669
GRIN2B	G00167_012	CCCTCTTCTCTCTTCTCTGTC	GGTGGAGATAGTCTCTCTTGG	408	NM_000834	12	12	13720271	13719864
GRIN2B	G00167_013	GGAGCAATTTGCTGGGAG	TGCGCTGACAGGTGCTGAGG	609	NM_000834	13	12	13717658	13717050
GRIN2B	G00167_014	ACTGTGACCAACCCACTCTG	CTTCTTCAAGGCTGACAGG	584	NM_000834	13	12	13717158	13716575
GRIN2B	G00167_015	ATGACTTTAAGCGCAGCTAC	GGACTTGTGTTGCGAAAGG	584	NM_000834	13	12	13716693	13716110
GRIN2B	G00167_016	GACCTGCGAGAAGGAGAACG	CACCCTCCGTGAGTACAGC	608	NM_000834	13	12	13716247	13715640
GRIN2B	G00167_018	GTTAAGGTGGTACGGGAGG	AATCCACTCTGCCCACTTGC	640	NM_000834	13.5	12	13715700	13715061
GRIN2B	G00167_019	ATGTGCCAAGCAGAGGTTG	CAACCTTGGCAGCTGAGCG	640	NM_000834	13.6	12	13715186	13714547
GRIN2B	G00167_020	TGCGCAATTTGTTCAACCAAG	CTTACCGCAAGTTTTCAGAGC	694	NM_000834	13.7	12	13714701	13714008
GRIN2C	G00377_001	AATCAATGCTCACTGTCTGC	AGCTGTGGGACAGAGGCTG	282	NM_000835	4	17	72847924	72847641
GRIN2C	G00377_004	AGAGGGAGATTCTTGCCGCC	GCATGAGCATACACAGACAC	379	NM_000835	5	17	72846996	72846618
GRIN2C	G00377_005	GTCCCATGCTGGGGTGTG	TCATCATTTGTGACAGCCGC	395	NM_000835	6	17	72846626	72846232
GRIN2C	G00377_006	CAGCTGGGCTGGGCTGAGAAG	CCCAAGCTGTACACACCTTC	324	NM_000835	7	17	72846171	72845848
GRIN2C	G00377_007	CTGTCTCAACCAGTGCTTTC	TCCCAAGGTTATCTCTGAAGG	296	NM_000835	8	17	72844192	72843897
GRIN2C	G00377_008	CAGTGGGGTGGGCTGAG	TCCCCAGACTCTGGGCTGACC	400	NM_000835	9	17	72843754	72843355
GRIN2C	G00377_009	GCTCTGGGGTGGGGTTTG	CCCAAGCTGAGAGCTAAGGCG	315	NM_000835	10	17	72843138	72842824
GRIN2C	G00377_010	TGTTTATAGACCTGTCTACG	CTTGGGAAGTCTGAGTGTGCG	358	NM_000835	11	17	72842477	72842120
GRIN2C	G00377_014	TTGTTGGTGTGCTTGTGAGG	AGGTTGTGAGCTGTGAGGCC	433	BC041128	12	17	72840721	72840289
GRIN2C	G00377_018	GCATGGGAGATGCTTCAAG	GAGGCCGAGAGGTGCCC	609	NM_000835	13.1	17	72839831	72839223
GRIN2C	G00377_019	CTGGGGACCCGACAGC	GTAGATCGGCAAGCACTATCG	508	NM_000835	13.2	17	72839399	72838892
GRIN2C	G00377_013	CTTCCGCGCCAGCTCTG	GAAGCCAGAAAAGCCCAATC	555	NM_000835	13.3	17	72839037	72838483
GRIN2C	G00377_015	CATCTGCTCTCTCCAAACCC	GAAGCAAGGATCTGGGCTGAC	402	BC059384	2.1	17	72851305	72850904
GRIN2C	G00377_016	CTCAGCAAGTTGGGCTCAAC	TCAGCCAGAAATCTCTCTCTC	400	BC059384	2.2	17	72851039	72850640
GRIN2C	G00377_017	GAGGGCTCCCGAATG	TCTGGTGAGCACTCTGTC	570	NM_000835	3.1	17	72848881	72848312
GRIN2C	G00377_002	GACGTGGTACGCTGTGAG	AGACAGGTGGGGTAAATGTTG	543	NM_000835	3.2	17	72848561	72848051
GRIN2D	G00354_004	GGACTTTCTCTGCTTCCCG	CTCTCCCTCTTATGCTCACG	284	NM_000836	4	19	48917175	48917458
GRIN2D	G00354_005	AGAGGATCATGAGGCGAGG	GTCCCAAGGGGAGGAGG	376	NM_000836	5	19	48917558	48917933
GRIN2D	G00354_006	TCTTTAGCTCTGAGTTTCCC	GTTTTGCCCAACCTGCTCTC	461	NM_000836	6	19	48917985	48918445
GRIN2D	G00354_007	CGTCTCTGCTCGCCGTC	AACTCCAGGTTTCTGATCTCC	320	NM_000836	7	19	48917981	48919500
GRIN2D	G00354_008	AGCGCTCAAGTGATCTTCC	GGGCTGCTGGGAATTTAG	331	NM_000836	8	19	48922363	48922691
GRIN2D	G00354_009	GGGCACTGGGAATTGTAG	CCEAAGATCCCTCTTAATC	415	NM_000836	9	19	48922747	48923161
GRIN2D	G00354_010	GGTGTGCTCTAGAATGGGTG	CTTGGCTATCTTGGAGCATC	310	NM_000836	10	19	48924964	48925273
GRIN2D	G00354_011	ACAGGTGAGGAGGTTCCAGG	TCTTTGCTCTTCTCTCTCCC	347	NM_000836	11	19	48944947	48945293
GRIN2D	G00354_012	GGGTCATAGGCTCTACATTG	ATCTCTTGTCTCTGCTCTCGG	398	NM_000836	12	19	48945322	48945719
GRIN2D	G00354_018	AAGAAGGGGCAAGAGCCCG	ACTATCCACATCTCCAGCCAC	463	NM_000836	12	19	48945278	48945740
GRIN2D	G00354_013	ACAGCGGGGTAAATGAGAG	TCGGTGGCTCTCTGTC	477	NM_000836	13.1	19	48945760	48946236
GRIN2D	G00354_031	TGCCCTTGGCTCTGAGGAGAG	CAGGCTGGGCTGAGTCCGACAGT	423	NM_000836	13.3	19	48946319	48946741
GRIN2D	G00354_028	CGCCGCGCCCAACACACAG	AAGGCCCGCCCGCTGTTCTCA	502	NM_000836	13.4	19	48946918	48947419
GRIN2D	G00354_002	CCATATTGGAGCTGTGTC	ACCATGAACACAGCTGAGCC	448	NM_000836	3.1	19	48907902	48908349
GRIN2D	G00354_003	AGATCCGCCCTGCTCTTCTG	CCCATTGTGAGCTCACTCATC	447	NM_000836	3.2	19	48908250	48908696
GRIN3A	G00226_001	GCAGGAGCTCTCAGGAC	AGGTTGTCTCCAGGCAATAG	483	NM_133445	1	9	104500340	104499858
GRIN3A	G00226_002	GGTTGGGGAGCACCTTG	ACCAAGCCGTTTCTCTCTC	494	NM_133445	1	9	104499972	104499479
GRIN3A	G00226_003	TATGGAGATTTTATCCAGAAC	GAATCTCCACACCACTAC	512	NM_133445	2	9	10449617	10449106
GRIN3A	G00226_004	TTAAGAACAGCACACCCACAG	CAGAGTCAAGATTCTGCAAGGC	447	NM_133445	2	9	10449220	104448783
GRIN3A	G00226_005	CAAGCTGAACTGCTCTTCC	GAATTTAATGGGCACTGATCATC	485	NM_133445	3	9	104433462	104432978
GRIN3A	G00226_006	CTGCTGGCCCACTCTG	GAGTGAGGAAGACGGCAGTG	469	NM_133445	3	9	104433082	104432614
GRIN3A	G00226_007	TTCCAGAGCCTTTTCTTCTC	GGCTGAGAGAAATATCTCTGTTGAT	538	NM_133445	4	9	104432764	104432227
GRIN3A	G00226_008	TACATGTTTCTGCGTGGGG	CAACGTGGGTATATTGACTGTG	390	NM_133445	3	9	104390814	104390425
GRIN3A	G00226_009	GTGCTGAGCTTCTATGCTTC	GTCTTGTGACTTTGCGATTTC	273	NM_133445	5	9	104385785	104385513
GRIN3A	G00226_010	CACAATGCTACTGTGGTGGC	GAGGCAACAGTATAGCATTC	308	NM_133445	6	9	104375865	104375766
GRIN3A	G00226_011	CCAAGAGTACAACAGATTGG	CATAGGCTCTGGATCTCTC	314	NM_133445	7	9	104341714	104341401
GRIN3A	G00226_012	AGAGGCCAAGGTTTGAACG	TGATTTCTTTGATAAATGCC	295	NM_133445	8	9	104340238	104339944
GRIN3A	G00226_013	TGGCTACACAGATGAAGACAGC	CCCCACACCTTTGTGATAG	510	NM_133445	9	9	104335868	104335359
GRIN3B	G00227_012	CTCAAGCAACCACTGGAGAG	CCCTGCTCTGAGACTCG	659	NM_138690	1.1	19	1000235	1000093
GRIN3B	G00227_013	CGATGAGTTTGTGCGGG	TCCTTCTCTCTGGGTTCTCGG	632	NM_138690	1.2	19	1000435	1001066
GRIN3B	G00227_002	CTGCTGGGGTGGGAGAG	TGCCCCACAGCTAGTG	429	NM_138690	2	19	10003060	1001448
GRIN3B	G00227_003	AGATGGACAGCTGGGAG	CTCTACAGTTGCTTTTGTG	471	NM_138690	2	19	10013315	1003805
GRIN3B	G00227_004	CTGGGACGGGTGAGAG	AGTAGCCGTGAGACACATTC	488	NM_138690	3	19	1004451	1004938
GRIN3B	G00227_005	AACACCACTTTTGTGTTTCC	GTACACGGGTGAGGAGAGCG	515	NM_138690	3	19	1004471	1005285
GRIN3B	G00227_006	ACTTACACAGCCCTTCTGTC	GTCCAGCTGACCACTCCC	514	NM_138690	3	19	1005128	1005641
GRIN3B	G00227_007	CTCGAGTGGCCAGGAC	CTTCCACCTCGAACCCTC	535	NM_1				

Discussion

Au cours de cette étude nous avons pu identifier plusieurs mutations potentiellement causales et/ou *de novo*. Le principal résultat de ce travail est l'association de mutations *de novo* dans *GRIN2A* et *GRIN2B* avec des cas sporadiques de SCZ et d'autisme, comme cela a été décrit récemment chez des patients atteints de déficience intellectuelle (Endele et al, 2010; Hamdan et al 2011). Bien que nous n'ayons aucune évidence fonctionnelle de l'effet de ces variants il convient d'indiquer qu'aucune autre mutation potentiellement causale n'a été identifiée dans *GRIN1*, *GRIN2A*, *GRIN2B* et *GRIN2D* alors que nous avons séquencé au moins 854 individus (incluant 283 volontaires sains) pour chacun de ces gènes laissant à supposer que ces gènes sont sous une pression de sélection forte et qu'ils sont critiques pour le développement et le fonctionnement cérébral normal. Ces résultats trouvent un support dans la mesure où la perte de fonction homozygote de *GRIN1* est létale chez l'animal.

Nous montrons aussi l'existence de mutations perte de fonction dans les gènes *GRIN2C*, *GRIN3A* et *GRIN3B*, à la fois chez des patients et des individus « contrôles ». Pour *GRIN3B* des mutations entraînant perte de fonction relativement fréquentes (10% d'homozygotes) ont déjà été rapporté dans la population générale (Niemann et al 2008) et nous confirmons cette observation. Pour *GRIN3A*, nous avons identifié trois mutations perte de fonction chez des individus contrôles et une chez un patient atteint de schizophrénie. Bien qu'il ne soit pas exclu que la mutation non-sens identifiée chez le patient puisse jouer un rôle dans l'étiologie de sa maladie via des mécanismes complexes d'épistasie ou d'interactions gène-environnement, nos analyses ne permettent pas de conclure à un tel mécanisme et il est vraisemblable que la perte de fonction hétérozygote de *GRIN3A* n'est pas causale en elle-même, puisque des sujets

contrôles en sont porteurs. Tout cela suggère des compensations fonctionnelles de la perte de fonction hétérozygote de *GRIN3A* et homozygote de *GRIN3B*.

Pour *GRIN2C*, il convient de rester plus mesuré. Nous avons identifié la même mutation non-sens chez un patient atteint de troubles du spectre autistique et chez un individu de la population générale. Bien que les mesures de fonctionnement global et de capacités d'apprentissage de cet individu semblent normales, son âge relativement jeune (7 ans) au moment de l'évaluation ne permet pas de fournir des conclusions fermes de l'effet de cette mutation sur le développement ultérieur de cet individu.

Enfin nous avons identifié un grand nombre de variants rares à la fois chez les patients et les contrôles mais nous ne pouvons pas conclure à un excès de mutations chez les patients par rapport aux individus contrôles

Chapitre 3 : Discussion générale

Alors que l'implication de facteurs génétiques dans la schizophrénie est clairement établie, la progression de la découverte des facteurs génétiques de la schizophrénie a été chaotique et s'est effectuée dans le sillage des avancées technologiques en génétique moléculaire. Les résultats générés par les dernières approches, permettant le séquençage à haut débit, amènent à réexaminer les différents modèles proposés pour expliquer la structure génétique de la maladie et de proposer des modèles alternatifs ou complémentaires aux modèles précédemment proposés.

Suivant le modèle des pathologies neurologiques, les premières études ont été des études de liaison, étudiant la transmission intrafamiliale, qui se sont révélées d'autant moins informatives que le mode de transmission est incertain. La seule conclusion consensuelle est que cette transmission n'est sûrement pas simple (Crow 2007).

Le domaine s'est ensuite tourné vers des études d'association, considérant l'implication d'un gène « candidat » particulier, choisi pour son rôle dans un processus physiologique possiblement en cause au vu des autres travaux du domaine ou pour sa localisation dans des zones préalablement repérées par des études de liaison.

Dans le sillage des avancées technologiques et en génétique moléculaire, identifiant une forte densité de polymorphismes répartis sur tout le génome (SNP), les chercheurs se sont ensuite tournés vers des études « génome entier », permettant, sans a priori particulier, d'étudier des polymorphismes répartis dans tout le génome, qu'ils soient ou non situés dans des gènes.

En réalité, toutes ces études (étude d'association gène candidat ou pangénomique) reposent sur l'hypothèse « variants fréquents-maladie fréquente ». Dans cette hypothèse le risque génétique

associé à la schizophrénie impliquerait des polymorphismes fréquents, chacun ayant un effet relativement faible pris individuellement, mais pouvant contribuer collectivement de façon significative à la maladie par des interactions gènesxgène ou gènesxenvironnements. Si ce modèle permet de rendre compte de la pénétrance incomplète de la maladie au sein des familles, il ne permet pas de rendre compte de toutes les situations, en particulier l'absence de tout antécédent familial chez certains patients. De plus, en considérant une hypothèse de transmission complexe et polygénique de variants fréquents, il apparaît que toute l'héritabilité de la schizophrénie ne pourra être expliquée par la coexistence de ces variants fréquents (Allen et al 2008; Jia et al 2010).

Cela conduit à envisager une autre hypothèse qui commence à recueillir des arguments depuis trois ans. Les études d'anomalies cytogénétiques avaient déjà attiré l'attention sur le fait que certains patients pourraient avoir des anomalies « rares » voire même uniques conférant un risque élevé à la maladie (Krebs 2009) et l'accumulation des résultats des études de CNV semblent confirmer l'implication de telles mutations rares et *de novo*. Cette hypothèse alternative, plausible et non exclusive à l'hypothèse des « variants fréquents » pourrait expliquer une part du chaînon manquant de la génétique de la schizophrénie. D'autre part, elle peut fournir des pistes pour la caractérisation des mécanismes physiopathologiques de la maladie, et pourrait permettre d'expliquer certaines observations épidémiologiques que ne permet pas d'expliquer l'hypothèse 'variants fréquents-maladie fréquente'. En effet, l'apparition de mutations de novo pourrait expliquer la relation entre âge parental élevé et le risque accru de maladies psychiatriques (Malaspina et al 2002), la persistance de la maladie dans la population générale malgré le manque de descendance (Bassett et al 1996) ou la

prévalence de cas sporadiques (patients n'ayant pas d'histoire familiale de troubles psychiatriques ou neuropsychiatriques).

Ces hypothèses, alternatives à l'hypothèse « variants fréquents-maladie fréquente », ont donc envisagé qu'une part substantielle des cas de schizophrénie pourraient être liée à des variants rares, chacun ayant un effet majeur conférant un risque élevé à la maladie. Ainsi, chaque cas pourrait être lié à un variant rare, différent d'une famille à une autre (mutation privée). Deux types de variants rares ou *de novo* ont été explorés et associés à la maladie, d'une part les anomalies chromosomiques ou CNV et d'autre part les mutations ponctuelles. Ces mutations pourraient toucher de multiples gènes différents sous la forme de multiples mutations rares différentes. Il a été montré que les CNV pourraient jouer un rôle dans l'étiologie de la schizophrénie en revanche bien que certaines mutations ponctuelles causales aient été identifiées dans l'autisme (Durand et al 2007; Gauthier et al 2011; Hamdan et al 2010; O'Roak et al 2011) ou le retard mental (Hamdan et al 2010; Hamdan et al 2009; Vissers et al 2010) ce champ d'investigation a été encore peu exploré pour la schizophrénie et fait l'objet de mes travaux.

Lorsque j'ai commencé mon doctorat en 2008 à Montréal, l'équipe du Dr Rouleau était déjà lancée dans un vaste projet de séquençage de gènes candidats chez des patients atteints de maladies neurodéveloppementales, générant des quantités considérables de données. Afin de faciliter l'analyse des résultats, nous avons focalisé notre intérêt sur un ensemble de gènes candidats fortement suspectés dans l'étiologie de la schizophrénie. Sur la base des études pharmacologiques, génétiques, et d'expressions, le système glutamatergique et les récepteurs NMDA, en particulier, nous sont apparus comme des excellents candidats.

Difficultés méthodologiques soulevées par l'étude des variants rares

La très grande quantité de données générées par le séquençage, même pour un nombre restreint de gènes, nous a contraint à nous focaliser sur un nombre restreint de variants et pose d'une manière générale le problème de l'analyse des données, de l'exome et du « genome sequencing ». Les premières études de séquençage à grande échelle ont permis l'identification des gènes responsables de maladies à transmission dominante. Ce fut le cas du gène *MYH3* dans le syndrome de Freeman-Sheldon (Ng et al 2009b). Cependant les stratégies d'analyse seront d'autant plus complexe que le caractère polygénique de la schizophrénie est établi.

Le choix qui a été fait est celui de se focaliser (1) sur les variants rares (présents chez moins de 5 individus sur l'ensemble des populations) (2) sur les variants *de novo* dans des familles où le cas de schizophrénie est sporadique, c'est à dire sans cas de schizophrénie chez les parents. La caractérisation des mutations *de novo* fournit une stratégie pour l'identification de mutations associées à la maladie. En extrapolant les résultats des études de séquençage à grande échelle qui ont inclus à la fois des individus et leurs parents (Conrad et al 2011; Durbin et al 2010; Roach et al 2010) on peut estimer qu'il y aurait, en moyenne, par individu, moins de deux mutations *de novo* dans la séquence codante et moins d'une mutation ayant un effet sur la séquence en acide aminé (Lynch 2010). Par conséquent la caractérisation d'une mutation peut être considérée comme un facteur de susceptibilité potentiel à la maladie d'autant plus que le taux de mutation *de novo* est augmenté dans la maladie (Awadalla et al 2010; Vissers et al 2010). Cette stratégie a déjà montré son efficacité pour l'analyse des données d'exome sequencing chez des patients atteints de troubles du spectre autistique (O'Roak et al 2011) et chez des patients atteints de schizophrénie (Girard et al, 2011)

Cependant cette stratégie pose certains problèmes. En premier lieu, la définition du caractère « sporadique » qui n'est pas consensuelle : généralement les cas sporadiques sont définis lorsqu'il n'y a pas d'apparentés atteints au 1^{er} degré (dans notre études les familles ont été retenues comme sporadique dès lors qu'aucun des parents n'était atteint). Par ailleurs, la qualité des données familiales recueillies est primordiale, car elles fournissent la base pour définir le diagnostic des apparentés au premier voire 2^d degré, sachant que certains phénotypes peuvent être incomplètement exprimés. Cette expression phénotypique incomplète pose également la question du choix stratégique de ne retenir que les mutations *de novo* et d'exclure les mutations transmises par l'un des parents, même rare, sur la simple base d'un diagnostic clinique, au mieux posé à l'issue d'un entretien structuré. On ne peut exclure que certaines mutations intéressantes aient été écartées alors que le parent, apparemment sain, était porteur de certaines caractéristiques endophénotypiques associées à la schizophrénie. Par ailleurs, ce type d'analyse nécessite également des cas cliniques extrêmement bien détaillés car on ne peut exclure que certains cas de schizophrénie sporadique soient des « phénocopies », expression d'altérations liées à des facteurs d'environnement en l'absence d'un terrain génétique particulier. Enfin, l'inclusion de patients d'ethnicités distinctes en particulier nous a posé de sérieux problèmes d'analyse, ces populations présentant un certain nombre de variants rares spécifiques.

Nous avons pu montrer l'association d'une mutation *de novo* tronquant la kinésine 17, impliquée dans le transport des récepteurs NMDA contenant la sous-unité NR2B. Nous avons aussi exploré les mutations rares et *de novo* dans les sous-unités des récepteurs NMDA et

montré l'association de mutation de novo dans *GRIN2A* et *GRIN2B* avec des cas de SCZ et d'autisme.

Nos résultats renforcent l'hypothèse d'un rôle du système glutamatergique, et des récepteurs NMDA en particulier, dans l'étiologie de la maladie. La kinésine 17 est une protéine motrice et est directement impliquée dans le transport le long des microtubules dendritiques de vésicules contenant la sous-unité NR2B (Setou et al 2000). Par conséquent la mutation tronquante observée chez le patient pourrait conduire à une diminution de l'expression synaptique des sous-unités NR2B comme observée *in vitro* (Guillaud et al 2003).

La sous-unité NR2B étant essentielle pour la localisation synaptique des récepteurs NMDA (Mori et al 1998), et ceux-ci jouant probablement un rôle dans la schizophrénie (Kantrowitz & Javitt 2010), il est vraisemblable que cette mutation joue un rôle dans l'apparition de la maladie chez ce patient. De façon intéressante une étude récente a montré dans un modèle animal de perte de fonction de la kinésine 17 que les déficits associés, à savoir des déficits d'apprentissage et de mémoire, sont principalement dûs à une diminution de l'expression synaptique des sous-unités NR2A et NR2B (Yin et al 2011).

L'identification de mutations de novo dans *GRIN2A* et *GRIN2B* avec des cas de SCZ et d'autisme dans nos travaux, confirme elle aussi l'implication des récepteurs NMDA dans l'étiologie de la maladie.

Nos résultats ainsi que ceux des études précédentes soulèvent au moins deux interrogations. La première d'entre elle est le rôle causal du gène dans la physiopathologie. Cette interprétation est rendue plausible (i) par le fait que ces mutations sont ponctuelles, touchant un seul gène par opposition aux CNV qui peuvent contenir plusieurs gènes ; (ii), en montrant l'absence de mutations entraînant potentiellement une modification fonctionnelle de la

protéine ou de son expression (site d'épissage, promoteur , etc) chez les individus contrôles et, pour les cas sporadiques, chez les parents (iii) et enfin en testant in vivo dans un modèle animal l'effet fonctionnel de ces mutations.

La deuxième interrogation concerne la spécificité des mutations aux phénotypes observés. Nous espérons limiter ce biais méthodologique de nos analyses, d'une part en reportant des mutations identifiées dans une cohorte de patients (ASD ou schizophrénie) mais absente dans les autres cohortes de patients (ASD, schizophrénie ou retard mental dans le cas de *KIF17*) et chez les contrôles et d'autre part par l'utilisation d'outils structurés standardisés et validés (entretiens diagnostiques en personne , données croisées familiales et par les praticiens etc...) pour la définition des diagnostics, permettant de limiter l'inclusion de phénotypes mal définis ou frontières. Le patient identifié avec la mutation tronquante *de novo* identifiée dans *KIF17* par exemple souffre incontestablement de schizophrénie et n'a pas de signe de retard mental. Notre stratégie n'excluait cependant pas non plus l'identification de mutations rares qui puissent être impliquées dans différentes pathologies comme ce fut le cas par exemple pour *SHANK3* avec la schizophrénie (Gauthier et al 2010) et les troubles du spectre autistique (Durand et al 2007; Gauthier et al 2009). Pour autant, il ne s'agit pas d'erreurs diagnostiques. Le détail des descriptions cliniques montre bien que les porteurs des mutations répondent au diagnostic de schizophrénie, et si deux des patients porteurs de la mutation dans *SHANK3* ont un retard mental et des traits autistiques associés, ceux ci sont absents des autres cas pour lesquels il n'y a pas de comorbidité et encore moins d'ambiguïté diagnostique. Notre stratégie permet de plus d'identifier des gènes dont la perte de fonction est susceptible d'être compensé fonctionnellement. L'identification de mutations tronquantes dans les sous-unités NR3 des

récepteurs NMDA, à la fois chez des patients et chez des individus contrôles laisse à penser que tel est le cas pour ces gènes.

Au final, nos résultats nous confortent dans l'hypothèse d'une implication glutamatergique, et des récepteurs NMDA en particulier dans l'étiologie de la schizophrénie et confirment qu'au moins une part des cas de schizophrénie pourrait être associée à l'implication de mutations rares à effet majeur. Cependant l'implication de ces mutations n'est pas exclusive du rôle des mutations fréquentes dans la schizophrénie.

D'une manière générale, les études génétiques entreprises pour les maladies complexes et pour la schizophrénie en particulier ont montré (i) que le génotype n'est pas prédictif du phénotype et (ii) que le même phénotype peut émerger de différents génotypes.

En ce qui concerne cette deuxième affirmation il est clair qu'il existe un très grand nombre de génotypes pouvant contribuer à un diagnostic de schizophrénie. Cependant, du fait de la très grande variabilité inter-individuelle, il est possible que les présentations cliniques dépassent les frontières nosologiques classiques et que la schizophrénie puisse représenter autant de maladies que de gène causal.

La translocation équilibrée $t(1;11)(q42.1;q14.3)$, identifiée pour la première fois dans une grande famille écossaise (St Clair et al 1990) est un bon exemple de la variabilité de l'expression phénotypique d'une même mutation. Alors que cette anomalie chromosomique est très bien définie d'un point de vue moléculaire (Millar et al 2000), elle est retrouvée associée à une gamme de présentation clinique incluant la schizophrénie, les troubles bipolaires ou la dépression majeure. Elle est aussi retrouvée chez des individus « sains ».

Les études génétiques menées sur d'autres maladies complexes fournissent une interprétation possible à cette diversité phénotypique. Dans le cas de la maladie de Huntington par exemple, le facteur de risque génétique de cette maladie est identifié, il s'agit d'une répétition de codon CAG dans le gène *IT15* ayant pour conséquence une répétition de glutamines dans la protéine huntingtine, Pourtant l'âge de début de la maladie est très variable pour un même nombre de répétitions. De façon intéressante, un certain nombre de modulateurs, génétiques et environnementaux, de la toxicité associée à la protéine ont pu être identifiés (Luthi-Carter et al 2010; Parker et al 2005).

Il se pourrait donc que l'expression phénotypique des mutations rares puisse être modulée par l'effet d'autres mutations sur d'autres gènes ou par l'exposition à des facteurs d'environnement, modulant par exemple la sévérité ou l'âge d'apparition de la maladie.

Dans ce cas, l'une des interprétations possibles du continuum entre autisme, schizophrénie et troubles bipolaires, pourrait provenir de cette modulation du phénotype associé à une mutation fortement pénétrante.

Par exemple, le phénotype pourrait être modulé par l'âge d'apparition des perturbations associées à une mutation causale. Si ces perturbations apparaissent au cours du développement précoce alors il semble probable qu'elles aient un impact plus important sur le développement normal, menant à des tableaux autistiques, alors que si elles apparaissent plus tard, à l'adolescence et lorsque les apprentissages principaux sont acquis, la même mutation pourrait conduire à des tableaux de schizophrénie.

Conclusion et perspectives

L'ensemble des résultats présentés dans ce travail sont en faveur d'une implication de variants rares et *de novo* dans l'étiologie de la schizophrénie et des troubles du spectre autistique. L'article d'Awadalla et al., présenté en annexe, a permis de montrer qu'il y avait un taux supérieur de mutations *de novo* ayant un effet potentiellement néfaste chez les patients. Celui de Piton et al. permet l'identification de plusieurs variants rares potentiellement associés à la schizophrénie et aux troubles du spectre autistique dans des gènes du chromosome X. L'apport de mes travaux (Tarabeux 2010) est de caractériser pour la première fois une mutation ponctuelle causale chez un patient atteint de schizophrénie et, d'impliquer le système glutamatergique, et les récepteurs NMDA en particulier, dans l'étiologie de cette maladie. (Tarabeux et al., 2011, en préparation)

Pris dans leur ensemble ces résultats confirment qu'au moins une fraction des cas de schizophrénie sont associés à des mutations rares ayant une forte pénétrance.

Tout cela suggère qu'il pourrait coexister différents modèles d'hérédité impliqués dans la génétique de la schizophrénie, allant de la transmission ou de l'apparition (pour les mutations *de novo*) de mutations monogéniques causales, à des modèles impliquant le rôle de variants fréquents, contribuant faiblement individuellement au risque associé à la maladie mais pouvant conférer un risque significatif collectivement, en passant par des modèles considérant à la fois le rôle de mutations rares, conférant un risque élevé, et de mutations fréquentes, et dans lequel l'expression pathologique serait dépendante des mécanismes complexes d'épistasie ou d'interactions gène-environnement.

Les stratégies envisageables pour le futur afin de mieux caractériser les facteurs génétiques de la schizophrénie devraient consister à considérer le séquençage à plus grande échelle (exome ou genome sequencing), de patients répondants aux critères les plus stricts de schizophrénie. Une telle approche optimiserait l'identification de mutations potentiellement causales, donc de processus physiopathologiques, d'une part en s'affranchissant des limites d'une approche gène candidat. Elle permettrait également d'avoir la cartographie génétique complète pour un individu. Néanmoins, une des limitations à cette approche sera la complexité des données qu'elle générera et la nécessité de développer des modèles d'analyse pour prendre en compte à la fois les mutations ayant un retentissement fonctionnel et les variants modulateurs. Par conséquent, il peut être raisonnable de continuer à inclure des patients « sporadiques » tels que définis dans notre approche ainsi que leurs parents. Cela pourrait permettre d'identifier rapidement de nouvelles mutations de novo causales et de caractériser les mécanismes physiopathologiques. Alternativement il sera nécessaire de décrypter les interactions gène-gène et gène-environnement. Par exemple il serait intéressant d'inclure dans les projets de séquençage à grande échelle des patients possédant la même mutation rare (DISC1 par exemple) et de caractériser le rôle des modificateurs de l'expression du phénotype. Un autre challenge sera d'avoir une stratégie de génétique inverse, recherchant à caractériser les profils phénotypiques associés à une mutation particulière.

Les progrès dans le décryptage du génome et des interactions génétiques complexe ainsi que la caractérisation du rôle des facteurs environnementaux devraient permettre de mieux connaître les mécanismes physiopathologiques sous-jacents à la maladie et ouvrir la voie à une recherche thérapeutique renouvelée et personnalisée ciblant mieux les causes de la maladie.

Annexes

Article 3 : Direct measure of the de novo mutation rate in autism and schizophrenia cohort.

Direct Measure of the De Novo Mutation Rate in Autism and Schizophrenia Cohorts

Philip Awadalla,^{1,2,3,15,*} Julie Gauthier,^{3,15} Rachel A. Myers,^{1,7,15} Ferran Casals,¹ Fadi F. Hamdan,^{2,3} Alexander R. Griffing,⁷ Mélanie Côté,³ Edouard Henrion,³ Dan Spiegelman,³ Julien Tarabeux,³ Amélie Piton,³ Yan Yang,³ Adam Boyko,⁸ Carlos Bustamante,⁸ Lan Xiong,³ Judith L. Rapoport,⁹ Anjené M. Addington,⁹ J. Lynn E. DeLisi,¹⁰ Marie-Odile Krebs,¹¹ Ridha Joher,¹² Bruno Millet,¹¹ Éric Fombonne,¹³ Laurent Mottron,⁴ Martine Zilvermit,¹ Jon Keebler,^{1,7} Hussein Daoud,³ Claude Marineau,³ Marie-Hélène Roy-Gagnon,² Marie-Pierre Dubé,⁵ Adam Eyre-Walker,¹⁴ Pierre Drapeau,⁶ Eric A. Stone,⁷ Ronald G. Lafrenière,³ and Guy A. Rouleau^{1,2,3,*}

The role of de novo mutations (DNMs) in common diseases remains largely unknown. Nonetheless, the rate of de novo deleterious mutations and the strength of selection against de novo mutations are critical to understanding the genetic architecture of a disease. Discovery of high-impact DNMs requires substantial high-resolution interrogation of partial or complete genomes of families via resequencing. We hypothesized that deleterious DNMs may play a role in cases of autism spectrum disorders (ASD) and schizophrenia (SCZ), two etiologically heterogeneous disorders with significantly reduced reproductive fitness. We present a direct measure of the de novo mutation rate (μ) and selective constraints from DNMs estimated from a deep resequencing data set generated from a large cohort of ASD and SCZ cases ($n = 1/4$ 285) and population control individuals ($n = 1/4$ 285) with available parental DNA. A survey of ~430 Mb of DNA from 401 synapse-expressed genes across all cases and 25 Mb of DNA in controls found 28 candidate DNMs, 13 of which were cell line artifacts. Our calculated direct neutral mutation rate (1.36×10^{-8}) is similar to previous indirect estimates, but we observed a significant excess of potentially deleterious DNMs in ASD and SCZ individuals. Our results emphasize the importance of DNMs as genetic mechanisms in ASD and SCZ and the limitations of using DNA from archived cell lines to identify functional variants.

Introduction

The rate at which human genomes mutate is critical to understanding every aspect of medical, statistical, and evolutionary genomics. To date, human mutation rate estimates have been indirectly inferred from a human-chimpanzee divergence approach,¹ from the analysis of mutations causing human Mendelian diseases,^{2,3} or, more recently, from next-generation sequencing in one nuclear family.⁴ These data suggest that there will be ~2 de novo mutations (DNMs) in the genome-wide coding regions per zygote, so that such mutations may contribute to some common diseases. Indeed, DNMs in individuals with complex disorders could explain genetic factors that are not detectable through genome-wide association studies. Deep resequencing of trios or families (e.g., patients and their parents) suffering from various diseases holds the promise of discovering DNMs that potentially could have a significant impact on disease prevalence

and severity. Disease-causing mutations are more likely to involve selectively constrained positions in which mutations are likely to be less tolerated or may have a substantial impact on fitness. If DNMs contribute significantly to a disorder, then there should be more functional and potentially deleterious mutations (1) in cases versus control samples and (2) in functionally constrained sites versus nonfunctional, and thus unconstrained (neutral), sites within the same cohort.

The disruption of gene function by rare deleterious penetrant mutations could represent an important cause of neurodevelopmental disorders such as schizophrenia (SCZ, MIM 181500) and autism spectrum disorders (ASD, MIM 209850). In fact, deleterious DNMs may explain observations such as the high global incidences of ASD (~0.45%)⁵ and SCZ (~0.4%)⁶ despite extremely variable environmental factors and reduced reproductive fitness,⁷ as well as increased risk with increasing parental age.^{8,9} Indeed, recent studies report an excess of de novo copy

¹Department of Pediatrics, Université de Montréal, Montréal, Quebec H3T 1C5, Canada; ²Centre Hospital Université Sainte-Justine Research Centre, Université de Montréal, Montréal, Quebec H3C 1G7, Canada; ³Centre of Excellence in Neuromics of Université de Montréal, Centre Hospitalier de l'Université de Montréal and Department of Medicine, Université de Montréal, Montréal, Quebec H2L 2W5, Canada; ⁴Department of Psychiatry, Hôpital Rivière-des-Prairies, Université de Montréal, Montréal, Quebec H1E 1A4, Canada; ⁵Centre de Recherche Institut de Cardiologie de Montréal, Department of Pharmacology, Université de Montréal, Montréal, Quebec H1T 1C8, Canada; ⁶Groupe de Recherche sur le Système Nerveux Central, Department of Pathology and Cell Biology, Université de Montréal, Montréal, Quebec H3C 3J7, Canada; ⁷Bioinformatics Research Center, North Carolina State University, Raleigh, NC 27606, USA; ⁸Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA; ⁹Child Psychiatry Branch, National Institute of Mental Health, Bethesda, MD 20892, USA; ¹⁰Center for Advanced Brain Imaging, Nathan S. Kline Institute, Orangeburg, NY 10962, USA; ¹¹University Paris Descartes, INSERM, Laboratoire de Physiopathologie des Maladies Psychiatriques, Centre de Psychiatrie et Neurosciences, U894, Sainte-Anne Hospital, Paris 75014, France; ¹²Department of Psychiatry, McGill University and Douglas Hospital, Montréal, Quebec H3A 1A1, Canada; ¹³Department of Psychiatry, McGill University and Montreal Children's Hospital, Montréal, Quebec H3Z 1P2, Canada; ¹⁴Centre for the Study of Evolution, School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK

¹⁵These authors contributed equally to this work

Table 1. Clinical Information for ASD and SCZ Individuals in which DNMs Were Confirmed

Sample	Final Diagnosis	Sex	IQ	Clinical Information	Comorbidity	Familial History of Psychiatric Illness	Age of Father at Birth (yrs)
S00004	Autism disorder	M	NA	ASQ ¹ score = 23	None	None	30
S00015	Asperger syndrome	F	NA	No physical dimorphism. ADI-R ² scores: social = 17, communication = 13, behavior = 7	Moderate scoliosis, hypopigmented skin patch	None	27
S00036	Autism disorder	M	NA	ADI-R scores: social = 23, communication = 14, behavior = 4	None	None	31
S00044	Autism disorder	M	NA	ADI-R scores: social = 24, communication = 10, behavior = 6	Minor anomaly: skull broad and flat on posterior aspect	None	40
S00096	Autism disorder	M	NA	ASQ score = 19	None	None	38
S00161	Schizoaffective	F	67	Childhood onset, age of onset 11 yrs. Patient with normal growth, no dysmorphic feature, speech impairment, and poor academic and social performance. ASQ score = 1	None	Father has lifetime depression and compulsive behavior	NA
S00285	Schizoaffective	M	NA	Schizoaffective disorder with age of onset of 19 yrs	Mild mental retardation	Parents are unaffected, two brothers are diagnosed with atypical chronic psychosis	NA
S00215	Schizophrenia	M	NA	No dysmorphic feature, moderate to severe emotional withdrawal	None	None	41

The following abbreviations are used: M, male; F, female; IQ, intelligence quotient; NA, not available.

¹ ASQ: Autism Screening Questionnaire (score > 15 = ASD).

² ADI-R: Autism Diagnostic Interview-Revised (total cutoff score for the communication and language domain is 8 for verbal subjects and 7 for nonverbal subjects. For all subjects, the cut off for the social interaction domain is 10 and the cut off for restricted and repetitive behaviors is 3).

number variants (CNVs) in ASD and SCZ compared to controls.^{10–12} We hypothesized that sequencing of families with affected individuals will identify an excess of missense relative to silent de novo mutations and that these mutations are candidate causal mutations for ASD and SCZ. As part of the Synapse-to-Disease Project (S2D), we resequenced synaptic genes in ASD and SCZ cases and resequenced a subset of these genes in a group of population controls. Such a resequencing project will capture DNMs at greater resolution, with the potential to unambiguously identify missense or frameshift mutations not detectable via linkage, association, or CNV methods. To test our hypothesis, we examined variants identified by resequencing 401 genes in a cohort of 285 ASD and SCZ individuals and for a subset of 39 of these genes in 285 population control individuals. Our analyses demonstrate a neutral mutation rate similar to that already reported and an excess of de novo deleterious mutations associated with the disease cohorts.

Subjects and Methods

Diagnostic Screening and Selection of Patients

The cohort of patients used for the sequencing of candidate genes for discovery of DNMs included 142 unrelated ASD patients (122 males and 20 females) as previously described,¹³ 65% of which

had no family history of ASD or related neurological disorders. All patients were diagnosed by using the Diagnostic and Statistical Manual of Mental Disorders criteria (Table 1). Depending on the recruitment site, the Autism Diagnostic Interview-Revised or the Autism Diagnostic Observation Schedule was used. In addition, the Autism Screening Questionnaire was completed for all the subjects. We excluded patients with an estimated mental age of <18 months, a diagnosis of Rett syndrome, or Childhood Disintegrative Disorder, as well as patients with evidence of any other psychiatric and neurological conditions, including birth anoxia, rubella during pregnancy, fragile-X syndrome, encephalitis, phenylketonuria, tuberous sclerosis, Tourette syndrome, or West syndrome. The 143 SCZ subjects (95 males and 48 females) were collected from five different centers and included 28 cases of childhood-onset schizophrenia (COS) and 115 sporadic or familial cases (with unaffected parents) of adult-onset schizophrenia or schizoaffective disorder.^{14–17} Sixty percent of the SCZ subjects had no family history of schizophrenia or other related neurological disorders. They were evaluated by experienced investigators who used the Diagnostic Interview for Genetic Studies (DIGS 3.0)¹⁸ or Kiddie Schedule for Affective Disorders and Schizophrenia, as well as multidimensional neurological, psychological, psychiatric, and pharmacological assessments. Family history for psychiatric disorders was also collected by using the Family Interview for Genetic Studies (FIGS). All DIGS and FIGS results were reviewed by two or more psychiatrists for a final consensus diagnosis based on DSM-III-R or DSM-IV at each center. Exclusion criteria included patients with psychotic symptoms mainly caused by alcohol, drug abuse, or other clinical diagnosis including major

cytogenetic abnormalities. We selected patients for which blood DNA was available so that DNMs could be validated and for which DNA was available from both parents to test for inheritance of the variants. The population control cohort (150 males and 135 females) consisted of unrelated individuals collected for the Quebec Newborn Twin Study (QNTS),¹⁹ in which DNA samples were available from both parents and both twins (either monozygotic or dizygotic); however, only one sibling was chosen randomly for sequencing. All samples were collected through informed consent following approval of each of the studies by the respective institutional ethics review committees. Ethnic origins of the grandparents were self-reported by the parents of probands and population control subjects. The ASD cohort is composed of French Canadians (85), other European Caucasians (54), and non-Caucasians (3). The SCZ cohort is composed of European Caucasians (136) and Asians (7). The control population is composed of French Canadians (204), other European Caucasians (55), non-Caucasians (18), and individuals of mixed origin (8).

Selection of Candidate Genes

The list of genes screened in the S2D project was generated from a repertoire of approximately 5000 potential synaptic genes compiled from several synaptic lists from different synapse databases, including the Genes-to-Cognition (G2C) database (which includes an extensive list of postsynaptic genes)²⁰ and the Synapse database (SynDB, which uses Synapse Ontology algorithms to mine potential synaptic genes),²¹ and from an extensive list of synaptic vesicle genes.²² It also comprises genes identified through manual searches of PubMed that are either localized at the synapse or affect synapse-related functions (i.e., plasticity, axon or dendrite outgrowth, dendritic spine morphology, learning, and memory).

The sequencing data reported here were generated, using Sanger technology, from the screening of the coding regions and splice site junctions of 122 X-linked and 279 autosomal potentially synapse-related genes (see Table S1 available online) in 142 ASD and 143 SCZ subjects. The excess of X-linked genes is due to the S2D selection procedure, which sought to include all potentially synaptic X-linked genes because of their importance in neurodevelopmental diseases²³ and because ASD is more common in males than females.²⁴ The genes on the autosomes were largely ones that encode glutamate receptors (including NMDA receptors, AMPA receptors, kainate receptors, and metabotropic glutamate receptors), as well as genes that encode proteins that interact with them, in particular those complexed with the NMDAR, a majority of which were identified by large proteomic studies²⁵ and reported in the G2C database.²⁰ The autosomal gene list is composed of 203 genes from the glutamate receptor complex (23 known glutamate receptors and 180 of their synaptic interactors) and 73 genes implicated in synapse function and/or cognition, interaction with ASD genes, or because of their disruption in the context of small CNVs or balanced translocations in patients with ASD, SCZ, or mental retardation. An additional three genes were included for which mutations were reported to cause ASD, SCZ, or the related neurodevelopmental disease, nonsyndromic mental retardation, which is known to coexist with ASD. Mutations in which multiple reports have previously found associations with diseases that are not related to ASD or SCZ were eliminated. In addition, data generated from the resequencing of 39 of these 401 genes in 285 population control samples were used. The 39 genes were resequenced in controls after discovering a de novo or deleterious mutation in ASD or SCZ samples.

DNA Preparation, Sequencing, and Variant Identification

Genomic DNA was extracted from peripheral blood lymphocytes and/or lymphoblastoid cell lines with Puregene extraction kits (Gentra System). A panel of seven microsatellite markers was used to confirm parentage for all samples.²⁶ To overcome the issue of limited DNA material, we performed the gene screening on DNA isolated from an Epstein-Barr virus transformed lymphoblastoid cell line for most cases ($n = 224$). The rest were done on blood DNA ($n = 61$). The ASD cell line samples had been frozen or regrown a maximum of two times. For the SCZ cell lines, 59 DNA samples were acquired from Coriell and are available through a request to J.L.E.D. All unique variants (heterozygous in a single individual) detected during the screen were tested in the parents' DNA. All potential de novo variants (not seen in the parents) were reconfirmed by reamplifying the fragment and resequencing the proband blood DNA and both his/her parents with reverse and forward primers in order to eliminate PCR or sequencing artifact. All de novo variants identified originally from cell line DNA were retested in the subject DNA extracted from blood to rule out variations that could have occurred during production or growth of the lymphoblastoid cell line. An identity DNA test was performed to eliminate any cell lines and/or blood inconsistencies caused by sample identification errors or nonpaternity. Primers were designed with the Exon Primer program from the UCSC Genome Browser. PCR products were sequenced on one strand with Sanger technology, done at the Genome Quebec Innovation Centre on a 3730XL DNA Analyzer System. PolyPhred (v. 5.04), PolyScan (v. 3.0), and Mutation Surveyor (v. 3.10, SoftGenetics) were used for variant detection.

Estimation of Base Pairs Screened

To estimate the number of base pairs (bp) screened, we determined the amount of coding and noncoding (intronic, UTR) sequence screened for each gene based on our PCR amplicon designs (Table S2). Briefly, genomic intervals were calculated based on forward and reverse PCR primer sequences for each amplicon. Because the DNA sequence overlapping each PCR primer is not surveyed, these genomic intervals did not include those corresponding to each PCR primer. Furthermore, because the first ~30 bp of sequence from each sequence read are of low quality, we trimmed 30 bp from each genomic interval (usually at the forward primer end). Overlapping amplicons were merged to form a single genomic interval. Then each genomic interval was annotated as to coding and noncoding sequence for each gene with tools available on the refGene table from the UCSC Genome Browser.

We defined functional and nonfunctional sites sequenced as those in which a nucleotide change would or would not lead to an altered protein sequence, respectively. We estimated that 71.2% of coding sites were functional, whereas nonfunctional sites were estimated at 28.8% of the coding and 100% of the intronic sites.²⁷ The number of CpG sites was estimated as 2.8% for functional sites and 1% for nonfunctional sites.²⁸ The corrections for increased mutation rates in CpG regions was $10 \times$ the number of CpG sites, yielding an effective bases sequenced of $10 \times \text{CpG sites} + \text{non-CpG sites}$ for both functional and nonfunctional sites.

Evaluation of False-Negative Mutation Calls

A subset of 71 of our ASD samples was also genotyped by Affymetrix 500K arrays. We tested our ability to detect heterozygous variants by comparing our heterozygous calls from the resequencing

Table 2. De Novo Mutations Discovered by Resequencing 458,877,850 Nucleotides of DNA in ASD, SCZ, and QNTS Control Individuals

Gene	Sample and Ref.	Diagnosis	Mutation Type	Mutation Location	Chr.	Position	Nucleotide Change and Genomic Context	Amino Acid and/or Structural Change	MAPP p Value
SHANK3	S00004	Autism disorder	INDEL	CODING	22	49500342	CGAGATTAGC(G/-)TAAGGGCCAC	Splice site delG	–
IL1RAPL1	S00015	Asperger syndrome	INDEL	CODING	X	29869731	CTTGGTGCTA(TACTCTT/-)GCTGCTTGTA	I367SfsX6	–
GSN	S00099	Asperger syndrome	INTRONIC	INTRONIC	9	123104277	GTGAGGCTGG(C/G)CCTGCCAGC	Within intron	–
KLC2	S00036	Autism disorder	MISSENSE	CODING	11	65788196	TACTATCGGC(G/C)GGCACTGGAG	R349P	0.001
KIF5C	S00044	Autism disorder	MISSENSE	CODING	2	149575030	GGACCGTAAG(C/T)GCTACCAGCA	R802C, R872C	0.001
FLJ16237	S00096	Autism disorder	MISSENSE	CODING	7	15393678	CCATCACTTA(T/C)TTTCCATATG	F279L	0.472
NRXN1	S02959	Schizophrenia	INDEL	CODING	2	50002821	CAGCACACGG(-/ACGG)GTATGGTCGT	G1402DfsX29	–
MAP2K1	S00237	Schizophrenia	INTRONIC	INTRONIC	15	64561310	CTTCTGTAC(G/T)GTCAGGGAGA	Within intron	–
SHANK3	S00161	Childhood-onset Schizophrenia	MISSENSE	CODING	22	49484091	GCATGACACA(C/T)GGCCTGGTGA	R536W	0.051
GRIN2B	S05650	Paranoid Schizophrenia	MISSENSE	CODING	12	13611351	CTTCTACATG(T/G)TGGGGGCGGC	L825V	<0.001
SHANK3	S00285	Schizoaffective, mild mental retardation	NONSENSE	CODING	22	49506476	TGCCCAGAG(C/T)GAGCTCTGGC	R1117X	–
KIF17	S00215	Schizophrenia	NONSENSE	CODING	1	20886681	GGAGCAGATA(C/A)TTCCTGGATG	Y575X	–
BSN	S00237	Schizophrenia	SILENT	CODING	3	49666988	GCACTGCAGT(G/C)GTAGACCTCC	V1665V	–
ATP2B4	S00182	Disorganized Schizophrenia	SILENT	CODING	1	201935404	TCATCCGAAA(C/T)GGTCAACTCA	N195N	–
SHANK3	S04261	QNTS, unknown	MISSENSE	CODING	22	49507364	GCCACCAGTG(C/T)CTCCCAAGCC	P1429S	0.107

screen of these 71 samples with calls made for overlapping genotyped SNPs on the array (Table S3). Of the 1649 heterozygous calls made in 90 autosomal SNPs on the Affymetrix 500K chip overlapping our screened amplicons, we failed to detect 41 of those heterozygous calls in our resequencing survey, suggesting that our false-negative rate is ~2%.

Prediction of Missense Severity

The potential consequence of each missense variant was evaluated with the MAPP,²⁹ PolyPhen,³⁰ SIFT,³¹ and PANTHER³² programs. Orthologous protein sequence alignments were obtained with tools available on the Galaxy Browser website for the generation of MAPP scores.

Statistical Analysis

Excess of functional relative to nonfunctional DNMs in each category (initial, CpG, non-CpG, and effective bases sequenced) was measured via (1) binomial test ($P[X \geq \text{number of functional DNMs} \mid q, \text{functional bases sequenced}]$, where q = neutral mutation rate) and (2) Fisher's exact test (FET). Functional DNMs are defined as missense and nonsense DNMs, whereas nonfunctional DNMs include silent and intronic DNMs.

Results

Identification of De Novo Mutations

By resequencing the coding and splice junction regions of 401 genes in 142 ASD and 143 SCZ samples (and 19 of these genes in 285 Quebec Newborn Twin Study [QNTS] samples), we identified 6184 DNA variants. Of these, 2437 were unique (i.e., heterozygous in 1 of 285 unrelated individuals tested). Each of these 2437 unique variants was resequenced in the proband and both parental samples to determine inheritance mode (transmitted versus de novo). A total of 15 unique variants was confirmed in the proband blood DNA sample but not detected in either parents' blood DNA sample (Table 2). These 15 validated DNMs are either germline-derived mutations or arose as somatic mutations in blood tissue. A further 13 variants were not detected in parents' DNA, nor were they detected in the proband blood DNA sample, and they were assumed to be generated during lymphoblastoid cell line development (Table 3).

Of the 15 confirmed DNMs, 14 were detected in the ASD and SCZ cohorts (2 nonsense, 5 missense, 3 frameshifting

Table 3. Cell Line Mutations Not Observed in the Blood Sample of Patients

Gene	Sample	Status	Mutation Type	Mutation Location	Chr.	Position	Nucleotide Change and Genomic Context	Amino Acid Change	Cell Line Origin
WWC1	S00056	AUT	SILENT	INTRONIC	5	167788404	CAGAAGGAAC(G/A)GTCTGTGTGG	–	UMontreal
PLCB1	S00068	AUT	MISSENSE	CODING	20	8585862	GATTTCACCTC(C/T)AGAAGTGATC	P209L	UMontreal
PLXNB3	S00009	AUT	MISSENSE	CODING	X	152694010	GGTGACCTGG(C/T)GGCCCCATTAC	A1431V	UMontreal
DRP2	S00093	AUT	MISSENSE	CODING	X	100383370	AAGCAGGCGA(C/T)GGTGGCCAGT	T203M	UMontreal
PSMD10	S00016	AUT	SILENT	CODING	X	107217953	TTGCGGCTTC(T/A)GCTGGCCGGG	S82S	UMontreal
MCF2	S00204	SCZ	SILENT	INTRONIC	X	138512219	TACAGTAATT(A/C)TTCAAGTATT	–	Krebs
SLC7A3	S00218	SCZ	INDEL	INTRONIC	X	70064365	CAGGTCACTAT(-/A)CAAATGTTTG	InsA 3' of exon	UMontreal
CAMK2A	S00191	SCZ	MISSENSE	CODING	5	149598465	CGAGGATGAA(G/A)ACACCAAAGG	D342N, D353N	UMontreal
GRPR	S00264	SCZ	MISSENSE	CODING	X	16080316	TCCCGGAAGC(G/T)ACTTGCCAAG	R261L	DeLisi/Coriell
ADAM22	S00193	SCZ	MISSENSE	CODING	7	87660477	AAAGTGAACC(G/A)ACAAAGTGCC	R860Q, R889Q, R896Q	UMontreal
ARHGAP6	S00261	SCZ	MISSENSE	CODING	X	11592643	GAGAGTCTCG(G/A)CCCTCGCTTG	G76D	UMontreal
ADD2	S00067	SCZ	NONSENSE	CODING	2	70744118	AAAGAAATTC(C/T)GAACCCCTC	R404X, R710X	UMontreal
RPS6KA6	S00274	SCZ	SILENT	CODING	X	83206731	ATCAGCGGTA(T/C)ACTGCTGAAC	Y672Y	DeLisi/Coriell

indel, 2 silent, and 2 intronic DNMs) and 1 was a missense DNM found in the population control group (Table 2). Five of the 11 point mutations in cases were transitions, four were CpG mutations, and six were transversions. Eight of the 14 DNMs detected among ASD and SCZ samples were disruptive to translation or protein structure and/or function, including three frameshift and two nonsense DNMs, and three missense DNMs were predicted to significantly disrupt protein structure via the computational prediction method MAPP.³³ Significant MAPP scores reflect a potentially damaging amino acid change and predict deleterious consequences. Reassuringly, MAPP scores relate to allele frequencies as would be predicted by population genetics, with increasingly deleterious alleles tending toward lower frequencies (Figure S1). Three MAPP p values of DNMs in SCZ and ASD samples are significantly low (Table 2); two were found in kinesin-encoding genes (R349P in *KLC2* [MIM 611729] and R802C in *KIF5C* [MIM 604593]). One nonsense mutation (Y575X) was also found in a kinesin-encoding gene (*KIF17* [MIM 605037]). One nonsense and one splice site deletion were found within *SHANK3* (MIM 606230), and a frameshift mutation was found in each of *IL1RAPL1* (MIM 300206) and *NRXN1* (MIM 600565). We have also confirmed the damaging predicted functional impact of the five de novo missense mutations via three other prediction programs (PANTHER, SIFT, and PolyPhen) (data not shown). Only one DNM was discovered in an X-linked gene (*IL1RAPL1*).¹² More detailed description of these genes and their potential role in ASD and SCZ will be presented elsewhere (unpublished data).

Estimates of the Neutral Human Mutation Rate

To calculate human mutation rates, we estimated the total initial count of bases we resequenced in all cohorts as

458.8 Mb (Table 4; Subjects and Methods). This includes 230.6 Mb of protein-coding and 228.3 Mb of intronic sequence. In the ASD and SCZ cohorts, exonic material sequenced is 215,186,702 bases and intronic is 218,145,769 bases. The total number of replacement sites in the cases is 153,704,787 and the number of silent (synonymous and intronic) sites is 279,627,684. In the QNTS cohort, 15,422,960 bases were exonic and 10,122,419 bases were intronic. The number of replacement sites is 11,016,400 and the number of silent sites is 14,528,979.

We distinguished functional from nonfunctional sites based on the effect of a mutation on transcription or translation of the protein at a given position (see Subjects and Methods). When addressing whether there was an excess of functional DNMs³⁴ relative to nonfunctional or silent base pairs, we calculated the “effective bp count.”³⁵ Because mutations are 10 times more likely to occur at CpG sites than at non-CpG sites, we calculated the total number of CpG and non-CpG sites for both functional and nonfunctional sites²⁸ and calculated the effective bp count (non-CpG sites + 10 × CpG sites) to account for increased mutation rates in CpG sites (Table 4).

We calculated the neutral mutation rate by examining the number of DNMs found in nonfunctional sites in our ASD, SCZ, and QNTS samples. In total, we sequenced ~294 Mb of nonfunctional DNA, in which we observed four DNMs (Table 2; in genes *GSN* [MIM 137350], *MAP2K1* [MIM 176872], *BSN* [MIM 604020], and *ATP2B4* [MIM 604020]). Because these mutations are unlikely to be pathogenic (referred to here as “neutral”), they allowed us to directly estimate the rate of neutral point mutations. We estimated this to be 1.36×10^{-8} mutations per site per generation (95% Poisson confidence interval: 0.34×10^{-8} , 2.7×10^{-8}). These estimates of neutral mutation rates are

Table 4. Base Pairs and DNMs Surveyed among ASD and SCZ Trios with No Family History

Site and Mutation Class		Initial Count ^a	CpG ^b	Non-CpG	Effective Count ^c
Functional	Nonsynonymous bases	96,065,492	2,689,834	93,375,658	120,273,996
	Nonsynonymous DNMs	6	3	3	6
Neutral	Synonymous bases	61,481,915	1,721,494	59,760,421	76,975,357
	Intronic bases	218,145,769	2,181,458	215,964,311	237,778,888
	Silent (synonymous and intronic) DNMs	4	2	2	4
p value	One-tail binomial test ^d	0.003	0.161	0.031	0.008
	Fisher's exact test	0.022	0.4041	0.1067	0.032

^a All statistics are calculated with base pairs and DNMs calculated for only trios with unaffected families.

^b Estimated 2.8% of initial coding sites and 1% of initial intronic sites are CpG sites.

^c To account for increased mutation rates: (non-CpG sites) + (10 × CpG sites).

^d $P(x \geq \text{number of functional DNMs} \mid \text{nonfunctional rate})$.

similar to, and not significantly different from, the estimate of 2.5×10^{-8} derived from phylogenetic analyses^{1,2} and the intergeneration estimate of 1.1×10^{-8} derived from next-generation sequencing data.⁴

Excess of Functional DNMs in ASD and SCZ Cohorts

If DNMs cause sporadic cases of ASD and SCZ, then DNMs will be more common in functional than in nonfunctional sites in our disease cohorts. Based on an observation of four DNMs in 294 Mb of neutral (silent and intronic) DNA, we expect 1.3 DNMs in the 96 Mb of nonsynonymous DNA sites to be resequenced in the cases with no family history of disease (65% of ASD cases and 60% of SCZ cases). However, among trios without family histories (Table 1), we observed six nonsynonymous DNMs surveyed in individuals, representing a significant enrichment of nonsynonymous DNMs ($p = 0.003$ in one-tail binomial test; $p = 0.022$ FET; Table 4). This excess remains significant even when we take into account that CpG dinucleotides mutate faster than other sites and are more common in exons than introns ($p = 0.008$ in one-tail binomial test; $p = 0.032$ FET; see Table 4 and Subjects and Methods).

If DNMs cause disease, we also expect point mutations with larger effects to be more frequent than expected in the disease group. Among our five nonsynonymous DNMs in trios with no family history, two are nonsense mutations (ratio 1:2.5), similar to previous estimates³ for DNMs causing Mendelian diseases (1:3.9) that are cataloged in the Human Gene Mutation Database (HGMD). Also, the ratio of synonymous to missense DNMs in the ASD and SCZ cohort is similar to that observed for HGMD.³ Under a neutral model,³ we would expect a ratio of 1 nonsense to 19.7 missense DNMs³ when only point mutations are considered. In HGMD, the number of missense to nonsense DNMs was significantly higher than the neutral expectation. Using a binomial test, our observed number of missense to nonsense DNMs was also significantly higher than the neutral expectation ($p = 0.04$), suggesting that some of the mutations are predisposed to be pathogenic. All of these observations

suggest an excess of potentially disease-predisposing DNMs in the SCZ and ASD cohort. Taken together, these lines of evidence suggest that mutations with functional effects are overrepresented within the synapse genes sequenced in individuals showing sporadic ASD and SCZ.

Comparing DNMs and Segregating Variant Ratios

Functional and nonfunctional segregating variants provide an expectation of the proportion of functional and nonfunctional DNMs. We compared the ratio of functional and nonfunctional DNMs to the ratios of the same classes of segregating variants in the ASD and SCZ cohorts (Table 5). Similar observations were found for the QNTS cohort. The comparison was significant when the functional and nonfunctional DNMs were compared to all segregating sites (FET, $p < 0.001$) and to unique SNP classes in the ASD and SCZ cohort (FET, $p = 0.003$). Given that the ratio of functional to nonfunctional was two times higher for DNMs relative to segregating sites, this suggests an excess of deleterious DNMs. Furthermore, given that rare SNP classes are likely enriched for slightly deleterious missense mutations,³ this significant comparison can be considered conservative. Under the expectation that most highly deleterious mutations will be selectively removed in one generation, the unique comparisons above

Table 5. Comparisons of Constraint for Genes Expressed at the Synapse in the ASD and SCZ Cohort

	Functional	Nonfunctional	Functional: Nonfunctional	p Value
Point and indel DNMs	10	4	2.5	–
Unique SNPs	785	1652	0.48	0.003 ^a
All SNPs	1306	4878	0.27	<0.001 ^a

Shown are the counts of point and indel mutations or segregating SNPs for the different categories of variation.

^a p value is the result of Fisher exact test comparisons for DNMs versus the two allele frequency classes of SNPs (unique or all).

provide insight into the proportion of deleterious mutations in humans that are selectively removed relative to segregating variation. Potentially disease-causing DNMs were more frequent than nonfunctional DNMs in our cohorts relative to expectations inferred from segregating mutations.

Discussion

In the present study, we have attempted to directly estimate the mutation rate with a large set of resequencing data generated from a common disease-based project. In addition, we have tried to validate that DNMs are a possible genetic factor of ASD and SCZ. There are three main conclusions that can be drawn from our study. First, the source of biological material (blood DNA versus cell line DNA) is crucial while doing experimental analyses with resequencing data seeking DNMs. All DNMs analyzed here were confirmed by resequencing, via standard Sanger technology, from DNA samples extracted from blood in the proband and in the parents. In so doing, we discovered that ~50% of our identified DNMs are the result of mutations that most likely occurred during the transformation and propagation of lymphoblastoid cell lines, thus creating false-positive DNMs. This observation was also recently stressed in CNV analyses by The Wellcome Trust Case Control Consortium.³⁶ This biological artifact, if unnoticed, would have significantly biased our results and would have contributed to a doubling of mutation rates for all classes of sites. Interestingly, 8 of 13 cell line mutations were X-linked, suggesting that this chromosome is particularly susceptible to the generation or accumulation of deleterious mutations after transformation of lymphoblasts with Epstein-Barr virus. These mutations, which are hemizygous in males, may also be positively selected because they contribute to higher fitness in cells carrying these mutations. An awareness of the high rate of mutation observed in cell lines, some of which are archived at the Coriell Institute, is critical to any large-scale whole-genome sequencing project, and potentially to those taking advantage of next-generation sequencing technologies, to capture rare and/or pathogenic mutations. Not only will the inherent error rate of the technologies be critical, but so, too, will the choice of samples and the way those samples are being maintained or cultured. Second, by using a direct calculation and classical Sanger sequencing, we validated the reported estimates of the neutral mutation rate in humans. Third, our study confirms the critical role that large-sample, high-resolution nucleotide surveys play in detecting potentially disease-causing DNMs.

We acknowledge that there are some weaknesses in our present study. The amount of resequencing in the controls is substantially lower than the resequencing in the cases. In fact, direct comparison in terms of sequences covered between cases and controls would be the optimal way to

directly estimate the rate of mutation and detect significant differences in mutation rate between the cases and controls. Nevertheless, our analyses show that the rate of potentially deleterious DNMs is significantly higher in functional compared to nonfunctional sites within the disease cohorts, suggesting a role of functional DNMs in the etiology of ASD and SCZ. Given that our estimate of the neutral human mutation rate is consistent with a recent genome-wide estimate⁴ and the accumulation of more direct observations of mutation, the confidence intervals of mutation rate estimates will begin to narrow. The rate of functional mutations in this survey is almost five times that of neutral or genome-wide rates, supporting our conclusions that we have likely detected mutations that are causal with respect to ASD and SCZ. By resequencing the genes in which DNMs were discovered among QNTS (random) participants, we were able to validate that these genes, among a substantial number of random individuals, do not carry functional DNMs, with the exception of one locus (*SHANK3*, Table 2). At *SHANK3*, a nonsynonymous mutation of low predicted functional impact was discovered in the QNTS cohort. *SHANK3* has been previously implicated in ASD³⁷ and may be a rapidly evolving gene in humans, with substantial neurological phenotypic impact.

By demonstrating that functional DNMs are at higher relative frequencies than segregating polymorphisms (Table 5), we show that DNMs may have a substantial role in ASD and SCZ etiology. The power of the genomics approaches employed here is that DNMs are not subject to the same demographic processes that shape segregating site variation. As a result, it is not necessary to test or correct for population structure, nor are our analyses subject to population stratification or admixture issues associated with GWAS analyses. By using a simple genomics approach that compares different classes of sites, we have sufficient power to map candidate mutations that are more likely to contribute to these diseases.

From sequencing only 8% of genes expressed in the synapse, functional DNMs were found in 5% of individuals with no family history, exhibiting a wide range of clinical phenotypes (see *Subjects and Methods* and Table 1). Although we biased our sampling strategy toward likely candidate genes, our predictions appear to have been poor regarding the X chromosome. It is therefore possible that by sequencing all 5000 synapse-related genes, we may uncover many of the mutations responsible for sporadic cases of ASD and SCZ. Furthermore, because nonfunctional DNMs are predicted to be relatively rare in ASD and SCZ genes (1.36×10^{-8} nonsynonymous DNMs per site), it may be easy to determine the likely causative mutations.

Supplemental Data

Supplemental Data include one figure and four tables and can be found with this article online at <http://www.cell.com/AJHG/>.

Acknowledgments

We would like to thank all the families and individuals who participated in this study. We are thankful for the efforts of the members of the Genome Québec Innovation Centre Sequencing and Bioinformatic groups. This work was supported by Genome Canada and Génome Québec and received cofunding from Université de Montréal for the Synapse-to-Disease (S2D) Project, funding from the Canadian Foundation for Innovation to both G.A.R. and P.A., and cofunding from the Ministère de l'Exploration, l'Innovation et l'Économie de Québec. G.A.R. holds the Canada Research Chair in Genetics of the Nervous System; P.A. holds career awards from the Fonds de Recherche Santé Québec and Génome Québec.

Received: January 22, 2010

Revised: July 22, 2010

Accepted: July 27, 2010

Published online: August 26, 2010

Web Resources

The URLs for data presented herein are as follows:

Coriell Institute for Medical Research, <http://www.coriell.org/>

Galaxy Browser, <http://main.g2.bx.psu.edu/>

Human Gene Mutation Database, <http://www.hgmd.cf.ac.uk/ac/index.php>

McGill University and Genome Québec Innovation Centre, <http://www.genomequebecplatforms.com/mcgill/>

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/>

PANTHER Classification System, <http://www.pantherdb.org/tools/csnpscoreform.jsp>

PolyPhen, <http://genetics.bwh.harvard.edu/pph/>

SIFT, <http://sift.jcvi.org/>

Synapse-to-Disease Project, <http://www.synapse2disease.ca/>

UCSC Genome Browser, <http://genome.ucsc.edu/>

References

1. Nachman, M.W., and Crowell, S.L. (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics* 156, 297–304.
2. Kondrashov, A.S. (2003). Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases. *Hum. Mutat.* 21, 12–27.
3. Kryukov, G.V., Pennacchio, L.A., and Sunyaev, S.R. (2007). Most rare missense alleles are deleterious in humans: Implications for complex disease and association studies. *Am. J. Hum. Genet.* 80, 727–739.
4. Roach, J.C., Glusman, G., Smit, A.F., Huff, C.D., Hubley, R., Shannon, P.T., Rowen, L., Pant, K.P., Goodman, N., Bamshad, M., et al. (2010). Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* 328, 636–639.
5. Rutter, M. (2005). Incidence of autism spectrum disorders: Changes over time and their meaning. *Acta Paediatr.* 94, 2–15.
6. Saha, S., Chant, D., Welham, J., and McGrath, J. (2005). A systematic review of the prevalence of schizophrenia. *PLoS Med.* 2, e141.
7. Bassett, A.S., Bury, A., Hodgkinson, K.A., and Honer, W.G. (1996). Reproductive fitness in familial schizophrenia. *Schizophr. Res.* 21, 151–160.
8. Croen, L.A., Najjar, D.V., Fireman, B., and Grether, J.K. (2007). Maternal and paternal age and risk of autism spectrum disorders. *Arch. Pediatr. Adolesc. Med.* 161, 334–340.
9. Malaspina, D., Brown, A., Goetz, D., Alia-Klein, N., Harkavy-Friedman, J., Harlap, S., and Fennig, S. (2002). Schizophrenia risk and paternal age: A potential role for de novo mutations in schizophrenia vulnerability genes. *CNS Spectr.* 7, 26–29.
10. Xu, B., Roos, J.L., Levy, S., van Rensburg, E.J., Gogos, J.A., and Karayiorgou, M. (2008). Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat. Genet.* 40, 880–885.
11. Stefansson, H., Rujescu, D., Cichon, S., Pietiläinen, O.P., Ingason, A., Steinberg, S., Fossdal, R., Sigurdsson, E., Sigmundsson, T., Buizer-Voskamp, J.E., et al; GROUP. (2008). Large recurrent microdeletions associated with schizophrenia. *Nature* 455, 232–236.
12. Piton, A., Michaud, J.L., Peng, H., Aradhya, S., Gauthier, J., Mottron, L., Champagne, N., Lafrenière, R.G., Hamdan, F.F., Joob, R., et al; S2D Team. (2008). Mutations in the calcium-related gene IL1RAPL1 are associated with autism. *Hum. Mol. Genet.* 17, 3965–3974.
13. Gauthier, J., Bonnel, A., St-Onge, J., Karamera, L., Laurent, S., Mottron, L., Fombonne, E., Joob, R., and Rouleau, G.A. (2005). NLGN3/NLGN4 gene mutations are not responsible for autism in the Quebec population. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* 132B, 74–75.
14. DeLisi, L.E., Shaw, S.H., Crow, T.J., Shields, G., Smith, A.B., Larach, V.W., Wellman, N., Loftus, J., Nanthakumar, B., Razi, K., et al. (2002). A genome-wide scan for linkage to chromosomal regions in 382 sibling pairs with schizophrenia or schizoaffective disorder. *Am. J. Psychiatry* 159, 803–812.
15. Gochman, P.A., Greenstein, D., Sporn, A., Gogtay, N., Nicolson, R., Keller, A., Lenane, M., Brookner, F., and Rapoport, J.L. (2004). Childhood onset schizophrenia: Familial neurocognitive measures. *Schizophr. Res.* 71, 43–47.
16. Joob, R., Rouleau, G.A., Lal, S., Dixon, M., O'Driscoll, G., Palmour, R., Annable, L., Bloom, D., Lalonde, P., Labelle, A., and Benkelfat, C. (2002). Neuropsychological impairments in neuroleptic-responder vs. -nonresponder schizophrenic patients and healthy volunteers. *Schizophr. Res.* 53, 229–238.
17. Mechri, A., Bourdel, M.C., Slama, H., Gourion, D., Gaha, L., and Krebs, M.O. (2009). Neurological soft signs in patients with schizophrenia and their unaffected siblings: Frequency and correlates in two ethnic and socioeconomic distinct populations. *Eur. Arch. Psychiatry Clin. Neurosci.* 259, 218–226.
18. Gourion, D., Goldberger, C., Bourdel, M.C., Bayle, F.J., Millet, B., Olie, J.P., and Krebs, M.O. (2003). Neurological soft-signs and minor physical anomalies in schizophrenia: Differential transmission within families. *Schizophr. Res.* 63, 181–187.
19. Lemelin, J.P., Boivin, M., Forget-Dubois, N., Dionne, G., Séguin, J.R., Brendgen, M., Vitaro, F., Tremblay, R.E., and Pélus, D. (2007). The genetic-environmental etiology of cognitive school readiness and later academic achievement in early childhood. *Child Dev.* 78, 1855–1869.
20. Croning, M.D., Marshall, M.C., McLaren, P., Armstrong, J.D., and Grant, S.G. (2009). G2Cdb: The Genes to Cognition database. *Nucleic Acids Res.* 37 (Database issue), D846–D851.
21. Zhang, W., Zhang, Y., Zheng, H., Zhang, C., Xiong, W., Olyarchuk, J.G., Walker, M., Xu, W., Zhao, M., Zhao, S., et al. (2007). SynDB: A Synapse protein DataBase based on synapse ontology. *Nucleic Acids Res.* 35 (Database issue), D737–D741.

22. Trinidad, J.C., Specht, C.G., Thalhammer, A., Schoepfer, R., and Burlingame, A.L. (2006). Comprehensive identification of phosphorylation sites in postsynaptic density preparations. *Mol. Cell. Proteomics* 5, 914–922.
23. Laumonnier, F., Cuthbert, P.C., and Grant, S.G. (2007). The role of neuronal complexes in human X-linked brain diseases. *Am. J. Hum. Genet.* 80, 205–220.
24. Skuse, D.H. (2000). Imprinting, the X-chromosome, and the male brain: Explaining sex differences in the liability to autism. *Pediatr. Res.* 47, 9–16.
25. Collins, M.O., Yu, L., Coba, M.P., Husi, H., Campuzano, I., Blackstock, W.P., Choudhary, J.S., and Grant, S.G. (2005). Proteomic analysis of in vivo phosphorylated synaptic proteins. *J. Biol. Chem.* 280, 5972–5982.
26. Gauthier, J., Champagne, N., Lafrenière, R.G., Xiong, L., Spiegelman, D., Brustein, E., Lapointe, M., Peng, H., Côté, M., Noreau, A., et al; S2D Team. (2010). De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for schizophrenia. *Proc. Natl. Acad. Sci. USA* 107, 7863–7868.
27. Eyre-Walker, A., and Keightley, P.D. (1999). High genomic deleterious mutation rates in hominids. *Nature* 397, 344–347.
28. Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci. USA* 103, 1412–1417.
29. Stone, E.A., and Sidow, A. (2005). Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. *Genome Res.* 15, 978–986.
30. Ramensky, V., Bork, P., and Sunyaev, S. (2002). Human non-synonymous SNPs: Server and survey. *Nucleic Acids Res.* 30, 3894–3900.
31. Ng, P.C., and Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 31, 3812–3814.
32. Thomas, P.D., Kejariwal, A., Campbell, M.J., Mi, H., Diemer, K., Guo, N., Ladunga, I., Ulitsky-Lazareva, B., Muruganujan, A., Rabkin, S., et al. (2003). PANTHER: A browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res.* 31, 334–341.
33. Stone, E.A., and Sidow, A. (2007). Constructing a meaningful evolutionary average at the phylogenetic center of mass. *BMC Bioinformatics* 8, 222.
34. Smith, N.G., and Eyre-Walker, A. (2001). Synonymous codon bias is not caused by mutation bias in G+C-rich genes in humans. *Mol. Biol. Evol.* 18, 982–986.
35. Eyre-Walker, A. (1998). Problems with parsimony in sequences of biased base composition. *J. Mol. Evol.* 47, 686–690.
36. Craddock, N., Hurles, M.E., Cardin, N., Pearson, R.D., Plagnol, V., Robson, S., Vukcevic, D., Barnes, C., Conrad, D.F., Gnanoulidou, E., et al; Wellcome Trust Case Control Consortium. (2010). Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* 464, 713–720.
37. Moessner, R., Marshall, C.R., Sutcliffe, J.S., Skaug, J., Pinto, D., Vincent, J., Zwaigenbaum, L., Fernandez, B., Roberts, W., Szatmari, P., and Scherer, S.W. (2007). Contribution of SHANK3 mutations to autism spectrum disorder. *Am. J. Hum. Genet.* 81, 1289–1297.

Article 4 : Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia.

ORIGINAL ARTICLE

Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia

A Piton¹, J Gauthier¹, FF Hamdan², RG Lafrenière¹, Y Yang¹, E Henrion¹, S Laurent¹, A Noreau¹, P Thibodeau¹, L Karemera¹, D Spiegelman¹, F Kuku¹, J Duguay¹, L Destroismaisons¹, P Jolivet¹, M Côté¹, K Lachapelle¹, O Diallo¹, A Raymond¹, C Marineau¹, N Champagne³, L Xiong¹, C Gaspar¹, J-B Rivière¹, J Tarabeux¹, P Cossette¹, M-O Krebs⁴, JL Rapoport⁵, A Addington⁵, LE DeLisi^{6,7}, L Mottron⁸, R Joober⁹, E Fombonne¹⁰, P Drapeau³ and GA Rouleau^{1,2}

¹Department of Medicine, Centre of Excellence in Neuromics, CHUM Research Centre, University of Montreal, Montreal, QC, Canada; ²CHU Sainte-Justine Research Center, Montreal, QC, Canada; ³Department of Pathology and Cell Biology and Groupe de recherche sur le système nerveux central, University of Montreal, Montreal, QC, Canada; ⁴INSERM U796, Physiopathologie des maladies psychiatriques, Université Paris Descartes and Centre hospitalier Sainte Anne, Paris, France; ⁵Child Psychiatry Branch, NIMH/NIH, Bethesda, MD, USA; ⁶VA Boston Healthcare System and Harvard Medical School, Brockton, MA, USA; ⁷The Department of Psychiatry, New York University Langone Medical Center, New York, NY, USA; ⁸Centre d'excellence en Troubles envahissants du développement de l'Université de Montréal (CETEDUM), Montreal, QC, Canada; ⁹Department of Psychiatry, Douglas Mental Health University Institute, McGill University, Montreal, QC, Canada and ¹⁰Department of Psychiatry, Montreal Children's Hospital, Montreal, QC, Canada

Autism spectrum disorder (ASD) and schizophrenia (SCZ) are two common neurodevelopmental syndromes that result from the combined effects of environmental and genetic factors. We set out to test the hypothesis that rare variants in many different genes, including de novo variants, could predispose to these conditions in a fraction of cases. In addition, for both disorders, males are either more significantly or more severely affected than females, which may be explained in part by X-linked genetic factors. Therefore, we directly sequenced 111 X-linked synaptic genes in individuals with ASD (n = 142; 122 males and 20 females) or SCZ (n = 143; 95 males and 48 females). We identified >200 non-synonymous variants, with an excess of rare damaging variants, which suggest the presence of disease-causing mutations. Truncating mutations in genes encoding the calcium-related protein IL1RAPL1 (already described in Piton et al. *Hum Mol Genet* 2008) and the monoamine degradation enzyme monoamine oxidase B were found in ASD and SCZ, respectively. Moreover, several promising non-synonymous rare variants were identified in genes encoding proteins involved in regulation of neurite outgrowth and other various synaptic functions (MECP2, TM4SF2/TSPAN7, PPP1R3F, PSMD10, MCF2, SLITRK2, GPRASP2, and OPHN1). Molecular Psychiatry advance online publication, 18 May 2010; doi:10.1038/mp.2010.54

Keywords: autism spectrum disorder; schizophrenia; X chromosome; synaptic genes; rare variants

Introduction

Autism spectrum disorder (ASD) and schizophrenia (SCZ) are two common neurodevelopmental disorders that typically appear during childhood, adolescence, or early adulthood. Twin and epidemiological studies strongly support the function of genetic factors in these diseases.^{1–3} Several linkage and association studies have been performed over the

past decades, with limited success.^{3–7} Positive linkage signals were obtained in numerous genomic regions and several genes or markers have been associated with ASD or SCZ; however, many of these results failed to be replicated in subsequent studies. The hypothesis implicating genetic heterogeneity, in which rare highly penetrant mutations (some of which may be de novo) in different genes specific to single families would predispose to these disorders, is supported by the recent findings that rare and de novo mutations in the NLGN4X, NLGN3, and SHANK3 synaptic genes cause ASD in a small number of families.^{8–12} Thus, direct resequencing of the complete coding regions of candidate genes is the best approach to identify such rare variants.

Correspondence: Dr GA Rouleau, Department of Medicine, Centre of Excellence in Neuromics, CHUM Research Centre, University of Montreal, 1560 Sherbrooke E, Room Y-3633, Montreal, Quebec, Canada H2L 4M1.

Received 16 October 2009; revised 10 April 2010; accepted 12 April 2010

A significant body of work suggests that ASD and SCZ are synaptic disorders. Reduced dendritic spines have been observed in neurons of these patients and nearly all the genes associated with these disorders are involved in the formation, regulation, or normal function of the synapse.^{13–19} Therefore, we screened our ASD and SCZ cohorts for variants in candidate genes that encode proteins involved in various synaptic processes, as part of the Synapse-to-Disease project (www.synapse2disease.ca). Here, we focused on X-linked synaptic genes, as several evidences have implicated the X chromosome in ASD and SCZ.^{20,21} The proportion of X-linked genes involved in brain development and cognition is high when compared with autosomal genes.^{22–24} Numerous X-linked genes have been implicated in non-syndromic mental retardation (NS-MR) and in syndromes associated with autistic features.^{23,25,26} In addition, linkage and association studies have identified several ASD and SCZ candidate regions on the X chromosome.^{27–36} Moreover, the prevalence of ASD or SCZ is increased in individuals with sex-chromosome aneuploidy, as is

the case in Turner and Klinefelter syndromes.^{23,37–39} Finally, the prevalence of ASD is four times higher in males than in females.⁴⁰ This sex-related difference in disease prevalence does not hold for SCZ; however, schizophrenic males seem to develop more severe symptoms at an earlier age than females.⁴¹

On the basis of the evidences presented above, we screened X-linked synaptic genes for rare damaging variants in a cohort of 142 ASD and 143 SCZ individuals (Figure 1a). The cohort of ASD includes patients with classic autism, Asperger Syndrome, and pervasive developmental disorder not otherwise specified, whereas the SCZ cohort includes patients with SCZ, schizoaffective disorder (SCZaff), and childhood onset schizophrenia (COS).

Materials and methods

Creation of the synaptic X-linked gene list

From a list of 1125 X-linked genes annotated in the NCBI Mapviewer Build 35.1, we performed a

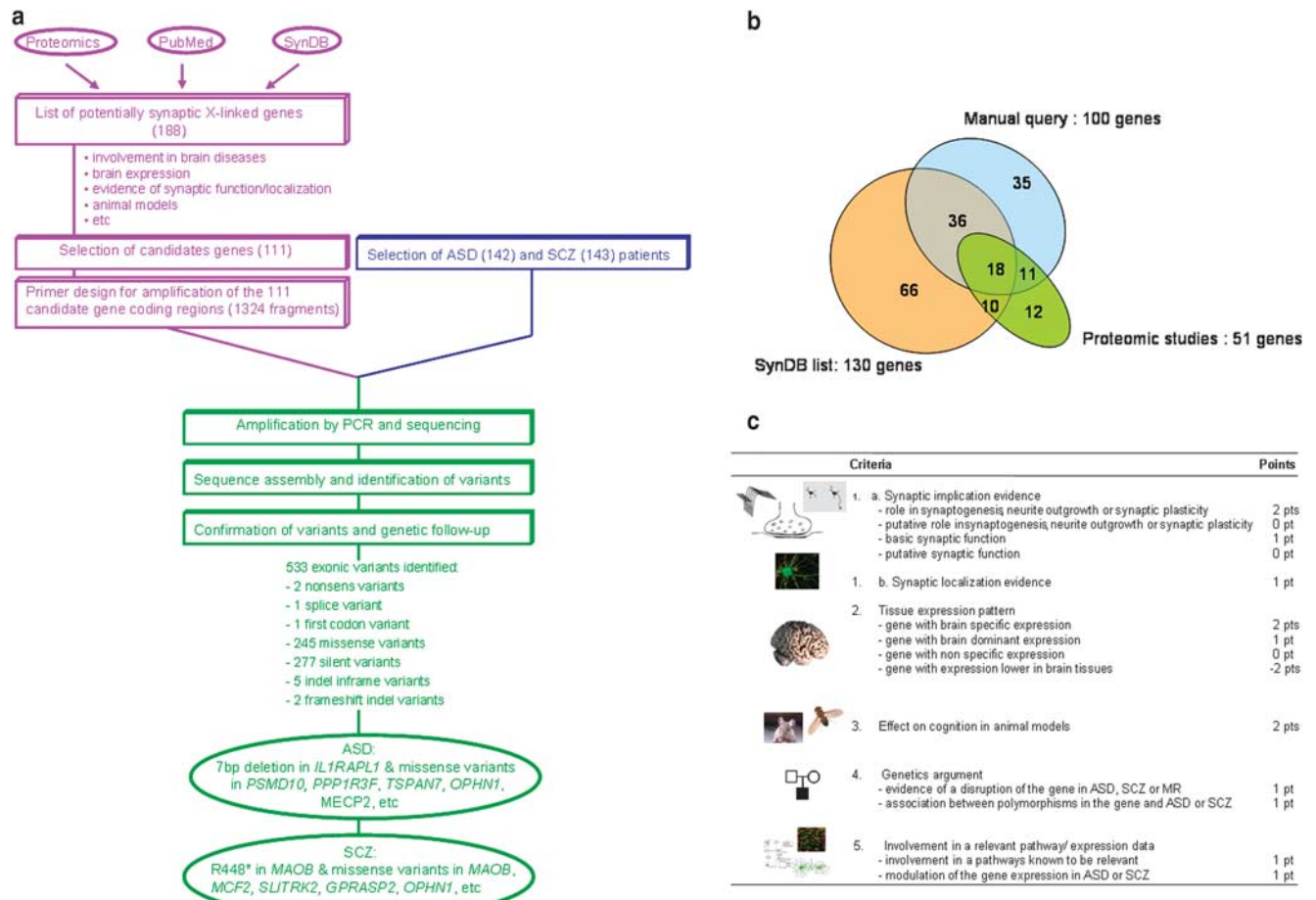


Figure 1 Strategy used for the selection and screening of synaptic candidate genes. **(a)** Schematic overview of the whole project: from the selection of candidate genes to the identification of genes probably involved in ASD and SCZ. **(b)** Representation of the number of synaptic genes identified according to the three different sources: proteomic studies, SynDB (Synapse Database), and manual query of PubMed (see Materials and methods; Supplementary Methods). The list of genes selected according to the different sources is provided in Supplementary Table 1. **(c)** List of the criteria used for the ranking and scoring of synaptic genes. The global score for each gene was the sum of all the points obtained using the various criteria. Further details are provided in Supplementary Tables 2 and 3.

quick search on each gene using information obtained from NCBI Gene and PubMed, as detailed in the Supplementary Information. Separately, we used different lists from proteomic studies^{42–44} and recovered 51 proteins encoded by X-linked genes. In parallel, we used a list extracted from Synapse DataBase (SynDB) (<http://syndb.cbi.pku.edu.cn/>), a new database dedicated to synaptic proteins that was created by the center for Bioinformatics of Peking University.⁴⁵ They found ~3000 human synaptic proteins, of which 130 are encoded by X-chromosome genes.

Selection of synaptic X-linked candidate genes

To select the best candidate genes, we built a scoring system based on different criteria. The first one was the involvement in a human disease: genes known to cause ASD, SCZ, or related syndromes (for example Rett syndrome RTT or NS-MR) were directly selected for the study, whereas genes known to cause other non-psychiatric diseases were excluded from the selection. Genes not clearly involved in any disease were ranked according to five different criteria relevant for ASD or SCZ: nature of synaptic function, evidence of synaptic localization, expression in brain tissues, cognitive impairment in animal models, genetic data, and involvement in a relevant pathway for ASD or SCZ. A score was calculated that corresponded to the sum of all the points attributed using the five criteria described above (Supplementary Table 3). Additional details for each criterion are provided in the Supplementary Information. We selected 113 synaptic genes for the variant screening; 25 based on their involvement in related diseases and 88 that had a score superior to zero. Two out of the 113 genes were not screened: *SYN1*, which had been sequenced and analyzed before the beginning of this project (Cossette *et al.*, unpublished data), and *NXF5*, because we were unable to design specific primers for this gene.

Patient and control cohorts

The 142 unrelated ASD patients (122 males and 20 females) were as described.⁴⁶ The 143 SCZ subjects (95 males and 48 females) were collected from five different centers and included 28 cases of COS and 115 sporadic or familial cases of adult onset SCZ or SCZaff. They were evaluated by experienced investigators using the Diagnostic Interview for Genetic Studies or Kiddie Schedule for Affective Disorders and Schizophrenia, and multidimensional neurological, psychological, psychiatric, and pharmacological assessments. Family history for psychiatric disorders was also collected using the Family Interview for Genetic Studies. All Diagnostic Interview for Genetic Studies and Family Interview for Genetic Studies have been reviewed by two or more psychiatrists for a final consensus diagnosis based on DSM-III-R or DSM-IV at each center. Exclusion criteria included patients with psychotic symptoms mainly caused by alcohol, drug abuse, or other clinical diagnosis including major cytogenetic abnormalities. Our population control group included 277 X chromosomes (from

103 males and 87 females) collected as part of the Quebec Newborn Twin Study. Ethnic origins of the grandparents were self-reported by the parents of probands and population control subjects. The ASD cohort is composed of French Canadians (85), of other European Caucasians (47), and of non-Caucasians (10). The SCZ cohort is composed of European Caucasians (133) and Pakistani or non-Caucasians (10). The control population is composed of French Canadians (129), of other European Caucasians (37), of South Americans (4), of non-Caucasians (3), and of individuals of mixed origin (16). Paternity and maternity of each individual of all families were confirmed using a panel of six highly informative unlinked microsatellite markers (D2S1327, D3S1043, D4S3351, D6S1043, D8S1179, D10S677). These markers were also used to confirm that cell line and blood samples corresponded to the same individual in the case of *de novo* mutations. An additional set of 285 ASD patients (247 males and 38 females) was used for the screening of *IL1RAPL1* and several exons of *TM4SF2/TSPAN7* and an additional set of 190 SCZ patients (118 males and 72 females) was used for the screening of monoamine oxidase B (*MAOB*) and several exons of *PLXNB3* and *GRPR*.

Gene resequencing and variant identification

Genomic DNA was extracted from blood samples or lymphoblastoid cell lines using the Puregene DNA kit (Gentra System, USA). The amplification of the coding regions of all isoforms of the 111 genes required the design of >1200 primer pairs, which was carried out using the Primers3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi), through Exonprimer (<http://ihg2.helmholtz-muenchen.de/ihg/ExonPrimer.html>), when possible. PCR reactions were performed in 384-well plates using 5 ng of genomic DNA, according to standard procedures. PCR products were sequenced at the McGill University and Genome Quebec Innovation Centre (Montreal, Canada) on a 3730XL DNA Analyzer. A fragment was considered successfully sequenced if the analysis of over 90% of the traces was possible. PolyPhred (v.5.04), PolySCAN (v.3.0), and Mutation Surveyor (v. 3.10, Soft Genetics) were used for variant detection analyses. In each case, unique novel exonic variants were confirmed manually by reamplifying the fragment and resequencing the proband and both parents using reverse and forward primers. For all *de novo* transmission variants, the specific exon was sequenced a second time using DNA extracted from blood instead of from cell lines, when possible. Variants found more than once or already described (in dbSNP <http://www.ncbi.nlm.nih.gov/projects/SNP/>, Exoseq <http://www.sanger.ac.uk/cgi-bin/humgen/exoseq/exoseqview> or in the literature) were considered as automatically confirmed (CA).

Genetic and statistical analyses

Rare variant follow-up. For each rare variant, the mode of transmission and the cosegregation between

the variant and the disease were evaluated by sequencing the corresponding exon in the parents and other available relatives. For the most relevant variants, we resequenced the corresponding exon in 190 individuals from our control population and/or in an additional 285 ASD and 190 SCZ patients.

Prediction of potential functional effect. The potential consequence of each missense variant was evaluated using the PolyPhen (<http://genetics.bwh.harvard.edu/pph/>), SIFT (<http://blocks.fhcrc.org/sift/SIFT.html>), and PANTHER (<http://www.pantherdb.org/tools/csnpscoreForm.jsp>) software.^{47–49}

Statistical analyses. The number of coding base pairs screened was calculated as follows: $3 \times \text{number of codons successfully sequenced} \times (\text{number of males} + 2 \times \text{number of females})$. The K_a of each gene was calculated as the number of non-synonymous (NS) variants identified divided by the number of NS sites screened. The K_s of each gene was calculated as the number of synonymous (S) variants identified divided by the number of S sites screened. K_a was also calculated for polymorphisms and for rare variants, ASD rare variants, or SCZ rare variants only. For the McDonald–Kreitman test, NS and S changes between human and macaque sequences were calculated using DNA^{SP} as described earlier by Tarpey *et al.*⁵⁰ The significance of differences between S and NS rare variants and human/macaque S and NS changes was evaluated for each gene using a Fisher's exact test. Significance was set at $P < 0.01$. The proportion of novel vs known (or rare variants vs polymorphisms) for S and NS variants was compared using a χ^2 test (1 df). The proportion of damaging vs benign predictions for rare variants

and polymorphisms was also analyzed using a χ^2 test (1 df).

Results

Selection and screening of candidate genes

From the 1125 X-linked annotated genes, we established a list of 188 genes encoding synaptic or potentially synaptic proteins using PubMed searches, published proteomic studies, and the SynDB (see Figure 1b; Material and methods; Supplementary Table 1). This corresponds to 17% of the total number of X-linked genes, with all the categories of synaptic functions being represented. We narrowed this list down to 111 genes by applying selective criteria and a prioritization scoring scheme (see Figure 1c; Material and methods; Supplementary Tables 2 and 3). To amplify the exons and the intronic boundaries of all available isoforms of each gene, we designed 1324 PCR amplicons and successfully optimized PCR conditions for 1263 fragments, which cover 96% of the coding regions. Fragments were amplified in the 285 patient samples and PCR products were sequenced (one strand).

We identified and confirmed 533 coding and splice variants (Table 1; Supplementary Table 4), which included 526 single-nucleotide changes and seven insertions/deletions (indels). A large number of silent (S) and missense (NS) variants as well as two nonsense variants, one splicing alteration, and one substitution of the first methionine were found. Five of the seven indels (in the *CACNA1F*, *PAK3*, *GPR50*, *GABRQ*, and *ATP6AP1* genes) were in frame and added/removed one or several amino acids. These indels did not target conserved amino acids predicted to be implicated in protein function, and, therefore, were unlikely to be pathogenic (Table 2). The two

Table 1 Distribution of the exonic variants, annotated by type of variant

Variant type	Total	Known vs novel		Pol vs rare	
<i>Indel</i>					
Indel frameshift ^a	2	0	2	0	2
Indel inframe	5	1	4	4	1
<i>Single nucleotide</i>					
Nonsense ^a	2	1	1	1	1
Splice ^a	1	0	1	0	1
First methionine ^a	1	0	1	0	1
Missense (NS)	245	60	185	86	159
Silent (S)	277	108	169	$P = 0.0004$	140
					137
					$P = 0.0004$

The 'known' variants were reported earlier in SNP public databases or in the literature, whereas other variants were 'novel'. The 'rare' variant group included all the novel variants identified in only one individual or in individuals sharing the same disease, either ASD or SCZ. In contrast, the 'pol' (polymorphism) group comprised all variants known as polymorphisms or that were not specific to one disease (variants found in both the ASD and SCZ cohorts). indel, insertion, or deletion of one or more nucleotides; single-nucleotide, single-nucleotide changes; NS, single-nucleotide non-synonymous missense variants; S, single-nucleotide synonymous variants; rare, variant found only once or in one cohort; pol, polymorphism, variant found in both cohorts or already described.

^aTruncating variants.

Table 2 List of the six truncating and five indel inframe variants identified during the screening of the 111 X-linked synaptic genes in 285 ASD and SCZ patients

Gene	Variant	Variant type	Already described	Disease	Sex	ASD	Frequencies				Comments
							SCZ	CONT	Inheritance	Familial info	
<i>MAOB</i>	c.1342C>T p.R448X	Nonsens	No	SCZ	M	0/142	1/332	0/190	Mother	1 SCZ brother does not have the variant	Removes the TM, IM and one part of the cytoplasmic domain.
<i>P2RY4</i>	c.1043G>A p.W348X	Nonsens	rs41310667	Both	M	2/130	2/140	3/190	n.a.	n.a.	Found in the three cohorts with similar frequencies. It might not affect protein function
<i>PCDH11X</i>	c.3034-1G>A	Splice	No	SCZ	F	0/139	1/113	0/190	n.a.	n.a.	Affects the end of the shortest isoform. Nonsense polymorphisms are described in dbSNP in the last exon, which is removed
<i>PLXNB3</i>	c.1A>G p.M1?	Unknown	No	SCZ	M	0/142	2/325	0/190	Mother	1 SCZ sister has the variant	This first methionine codon is unlikely to be the real initiator codon. It is not conserved between species and does not match the Kozak consensus sequence.
<i>HS6ST2</i>	c.23delT p.V8AlaX26	Indel (fishift)	No	SCZ	F	0/142	1/143	1/190	Mother	1 SCZ half-brother has the variant	Found in a girl from the CONT cohort. The variant is inherited from her unaffected father. It might be a polymorphism.
<i>IL1RAPL1</i>	c.1100-1106 delTACTCTT p.I367fsX6	Indel (fishift)	No	ASD	F	1/422	0/143	0/189	<i>de novo</i>	No familial history	Affects protein function (Piton <i>et al.</i>) ⁵¹
<i>CACNA1F</i>	c.2440 delGAGGAAGAA p.E814-816del	Indel (inf)	No	Both	Both	3/142	9/142	n.a.	n.a.	n.a.	Deletion/insertion of one or several Glu in a region with Glu repetition.
<i>PAK3</i>	c.525delGAT p.D175del	Indel (inf)	No	SCZ	M	n.a.	1/94	n.a.	n.a.	n.a.	Deletion of one Asp in a region with Glu repetition. It might not have a damaging effect on protein function.
<i>GABRQ</i>	c.1432delATT p.I478del	Indel (inf)	No	Both	M	1/141	1/143	n.a.	n.a.	n.a.	One polymorphism I478F was earlier described at this position. It might not be a critical amino acid for the protein.
<i>GPR50</i>	c.1505-1516 del12bp p.502-505del	Indel (inf)	No	Both	Both	57/142	66/143	n.a.	n.a.	n.a.	Already described with no association found with SCZ.
<i>ATP6AP1</i>	c.97insGCCGCG p.A33_A34insAA	Indel (inf)	No	Both	Both	2/142	1/143	n.a.	n.a.	n.a.	Insertion of 2 Ala found in the two cohorts. It might be a polymorphism.

Abbreviations: ASD, autism spectrum disorder; CONT, control population; F, female; fishift, frameshift; IM, intermembrane domain; inf, in frame; M, male; n.a., non available; SCZ, schizophrenia; TM, transmembrane domain.

other indels caused a frameshift and are described in the next section.

Variants that may cause protein truncation

In total, six potentially truncating variants, four single-nucleotide substitutions and two indels were identified (Table 2). The nonsense variant found in *P2RY4* (c.1043G>A, p.W348X) was clearly a polymorphism, referenced as rs41310667. The frameshift variant detected in *HS6ST2* (female (F): c.23delT, p.V8AfsX26), which was predicted to truncate a large portion of the protein, was also found in one control female and her father and was presumed to be a polymorphism. Variants affecting the consensus splice site (F: c.3034-1G>A) and initiation methionine (2 males (M): c.1A>G, 2 M) of *PCDH11X* and *PLXNB3*, respectively, were not predicted to have a strong damaging effect because of their relative position in the protein (see comments in Table 2). In contrast, the 7bp deletion (F: c.1100-1106del-TACTCTT) in the *IL1RAPL1* gene, which caused a frameshift p.I367fsX6 in one girl with Asperger Syndrome, led to a loss of function of *IL1RAPL1*, as described earlier.⁵¹ The screening of *IL1RAPL1* in additional ASD patients led to the identification of a new missense variant (M: 59A>G; p.K20R), which was not present in the control population. Finally, a nucleotide substitution (M: c.1342C>T) leading to a premature stop codon (p.R448X) was identified in one boy in the gene encoding *MAOB* (Figure 2a). This patient had a SCZaff disorder with an early age of onset and the R448X variant was inherited from his mother, who had a history of major depression. Moreover, there was some familial history of SCZ on the mother's side. The patient's brother, who had a

diagnosis of SCZ, did not carry this variant. However, the proband displayed a different diagnosis (SCZaff vs SCZ) and an earlier age of onset (15 vs 22 years) than his elder brother. This R448X variant was predicted to delete the last 73 amino acids of the MAOB protein, which encompass the mitochondrial anchor to the outer membrane and the intermembrane domains; this truncation could affect the targeting of the protein to mitochondria. Sequencing of 190 additional SCZ patients led to the identification of three novel missense variants, c.326C>T (M: p.P109L), c.958G>A (M+F: p.E320K), and c.1003G>A (M: p.G335S), in three boys, all transmitted by a heterozygous unaffected mother (Figure 2b; Supplementary Table 6). Moreover, the E320K variant was also found in one other girl with SCZ. These variants all affected evolutionarily conserved amino acids located in the cytoplasmic domain of the protein and may affect protein function, particularly the P109L variant, which modifies an amino acid involved in binding to the mitochondrial membrane.⁵² The exons of *MAOB* that contain the variants described were sequenced in the 190 control individuals and no NS changes were identified.

Analysis of variants identified gene per gene

Taking into consideration the different RefSeq isoforms of all the genes sequenced, in this study, we screened an estimated 71,723 codons per patient (Supplementary Table 5), which corresponds to a total of 76 Mb of coding sequence. S and NS variants were found in an approximately equal proportion, which is consistent with the findings of a recent analysis of the whole exome of one individual⁵³. For each gene, we represented the number of S + NS variants as a

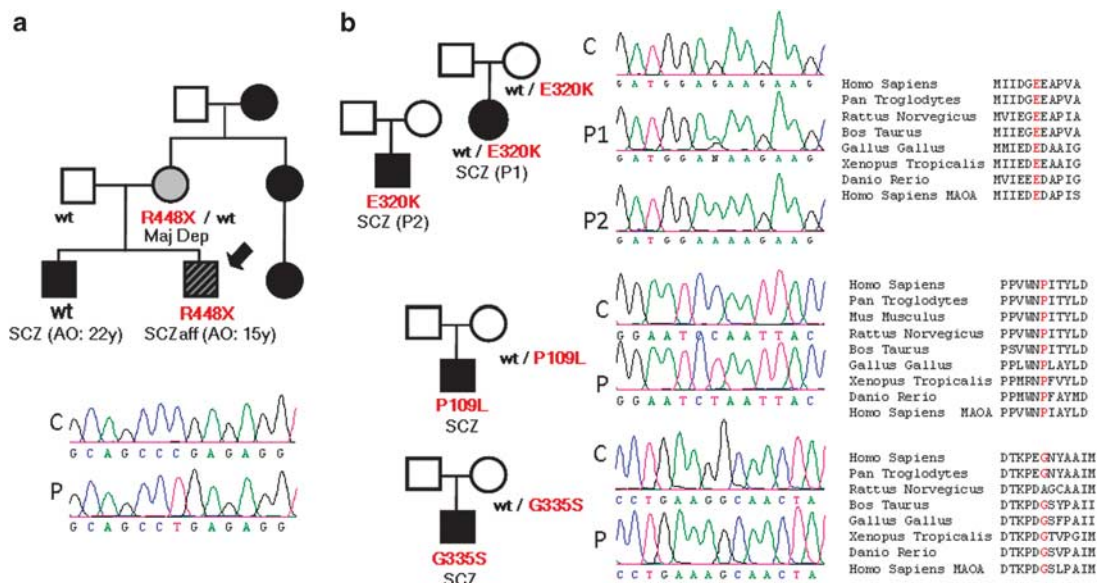


Figure 2 *MAOB* variants identified in SCZ individuals. The corresponding family pedigree (the proband is indicated by an arrow) and chromatograms for the proband (P) and control (C), as well as amino-acid conservation, are shown for each variant. (a) Nonsense variant identified during the first screening of 143 SCZ patients. (b) Missense variants identified during the screening of 190 additional SCZ patients.

function of the length of the coding sequence (Supplementary Table 5; Supplementary Figure 1a). Some genes (such as *USP9X*) had less variants than expected, according to their coding-sequence length, whereas others (for example *SHROOM2* or *HCFC1*) exhibited an accumulation of variants; however, we could not establish a correlation between these findings and the chromosomal location of the different genes (data not shown). Genes such as *HCFC1* had significantly less NS than S variants, which suggests that variation in their coding sequence is not well tolerated (Supplementary Figure 1b). In contrast, an excess of NS variants were observed in genes such as *PLXNB3*, *GPR50*, or *HS6ST2*. For each gene, we calculated the K_a , which is the number of NS variants divided by the number of NS sites. The analysis of genes with a high K_a in our cohorts did not allow the distinction between a highly polymorphic gene (in which greater coding-sequence variation is tolerated) from a gene enriched for disease-causing variants.

Excess of rare and damaging NS variants

We classified the variants identified during our study either as rare variants or polymorphisms. The 'rare variant' group included all novel (not reported in any database) variants identified in only one individual or in individuals sharing the same disease, either ASD or SCZ (as they could be rare variants shared by patients with a recent common ancestry). The 'polymorphism' group included all known variants or all variants not specific to one disease, which likely represent neutral polymorphisms.

Considering NS variants as more likely to have an effect on gene function, as they cause an amino-acid change, and S variants as a pool of generally neutral variants, we identified two parameters that were associated with an uneven distribution of NS and S variants. Only 30% of the variants identified were already known (either present in SNP databases (dbSNP and ExoSeq) or described in the literature). However, the distribution of novel vs known variants was significantly different between S and NS variants ($P=0.0004$), with an excess of NS variants being novel (Table 1). Perhaps more relevant, the grouping of variants into rare variants and polymorphisms revealed a significant difference in the distribution of S and NS variants ($P=0.0004$) toward the identification of more NS rare variants than expected. The analysis of the potential effect of NS missense variants using various prediction programs showed that, according to at least two out of the three algorithms tested, rare variants were significantly more damaging than polymorphisms (Supplementary Table 4; Figure 3), which suggests that they may include disease-causing mutations for ASD or SCZ. Although we decided to focus on NS changes in our study, we are aware that, in some cases, certain S variants may have an effect on mRNA splicing or stability.^{54,55}

Genes that present an excess of NS rare variants

Some genes had a higher K_a for rare variants ($K_{a_{rare}}$) than their K_a of polymorphisms ($K_{a_{pol}}$) (Supplemen-

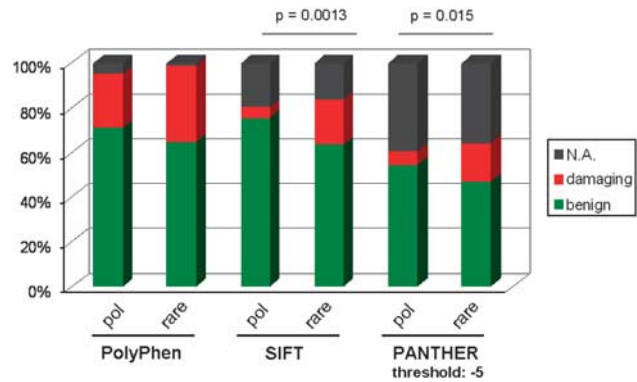


Figure 3 Potential effect on protein function of the different missense variants identified, according to three prediction programs: PolyPhen, PANTHER, and SIFT. The variants were divided into the 'rare variant' (rare) and 'polymorphism' (pol) groups. The percentage of variants associated with benign (PolyPhen) or tolerated (SIFT) predictions or with a score >-5 (PANTHER) are represented in green, whereas variants associated with a possibly or probably damaging effect (PolyPhen), affecting protein function prediction (SIFT), or with a score <-5 (PANTHER) are represented in red. The gray portion represents variants for which prediction was unknown (PolyPhen), given with low confidence (SIFT), or not available (PANTHER).

tary Table 5; Supplementary Figure 2). Among them, some had a higher $K_{a_{rare}}$ in one cohort than in the other; these included genes such as *TM4SF2/TSPAN7*, *HNRPH2*, *PSMD10*, and *CACNA1F* for ASD and *RBM3*, *PNCK*, *SLITRK2*, or *KCND1* for SCZ.

To identify genes that exhibited a greater number of rare NS variations in the patients than that expected from their rate of evolution, we performed a McDonald-Kreitman test (Supplementary Table 5). For the SCZ cohort, significant differences were obtained for *SLITRK2* ($P=0.0006$), a neuronal transmembrane protein family SLITRK, which affects neurite outgrowth.⁵⁶ We identified in this gene two novel missense variants c.265G>A and c.1646C>T (F: p.V89M and F: p.S549F) in girls with SCZ and in their affected siblings. Significant results were also obtained for *SEPT6* ($P=0.004$), and *KCND1* ($P=0.004$). Surprisingly, *MECP2*, the main Rett syndrome gene, showed an accumulation of NS rare variants ($P=0.004$) in the SCZ cohort. A significant excess of NS rare variants in the ASD cohort was obtained for *CACNA1F* ($P=0.0011$). This gene is involved in congenital stationary night blindness of type 2. A missense mutation (I745T) in *CACNA1F* was described in a large family with congenital stationary night blindness of type 2, in which five of the affected males had intellectual deficits and three of these five individuals were diagnosed with ASD.⁵⁷ Interestingly, other members of this calcium channel family have been linked to ASD (*CACNA1C*,⁵⁸ *CACNA1H*,⁵⁹ and *CACNA1G*⁶⁰). Although the two NS variants identified in two ASD patients (2 M: p.V635I and M+F: p.N746T) were also found in the control cohort, the four unique missense

variants (M: p.L977F, M: p.D1966N, M: p.Q1625R, and M: p.G1684R) were absent in the control population; however, the analysis of cosegregation and functional effect predictions did not support their implication in ASD. *CDKL5*, which causes Rett syndrome,⁶¹ and *PLXNB3*, a plexin acting as a receptor for Semaphorin 5A, also exhibited an accumulation of NS rare variants ($P=0.0068$ and $P=0.0012$, respectively), but this was not significant for ASD or SCZ separately. To deal with the population structure issue, we performed again the analysis excluding rare variants from patients with a non-European-Caucasian origin (10 ASD and 10 SCZ), and we obtained similar results.

Transmission, gender, and disease distribution of rare variants

To evaluate the mode of transmission of each rare variant, we reamplified the corresponding genomic fragments and resequenced them in the proband and his/her parents. As the large majority of parents were non-affected, we can discern modes of transmission that favor the implication of rare variants in disease, that is *de novo* inheritance or transmission from a heterozygous mother to a hemizygous boy. We need to be particularly cautious with the *de novo* variants, as we found that a non-negligible proportion of the *de novo* variants identified here were artifacts associated with the use of DNA extracted from lymphoblastoid cell lines and could not be confirmed in corresponding blood DNA. Among the eight *de novo* variants initially identified in this study, only the deletion in *IL1RAPL1* and the missense in *MECP2* were confirmed as *bona fide de novo* variants. Two of these variants were cell-line variants and four were not tested because blood DNA was not available, but they are likely to be cell-line artifacts, as for most of them, they were found as heterozygous variants in male patients.

On the basis of the number of chromosomes tested, we observed no significant differences in the number of NS rare variants found in males and in females. Only one NS variant (F: c.175G>A; p.D59N in the *TRO* gene) was homozygous in one girl from a Pakistani family. This variant was predicted to have a benign effect on protein function and did not segregate with SCZ in the family. However, some of the NS rare variants identified in heterozygous females could also be causative. Indeed, in the case of variants transmitted by the unaffected mother, it could be a difference in the X-inactivation patterns that cause a phenotype in some carrier females and not in others. However, we decided not to test X-inactivation pattern in a systematic way for all the rare NS variants identified in females, given that extrapolation of X-inactivation pattern results in lymphoblastoid cell lines DNA, or even in blood, might be wrong to estimate what occurs in brain tissues. We did not observe any significant difference in the number of NS rare variants identified between the ASD and the SCZ cohorts. However, among the

individuals of European-Caucasian origin, more SCZ than ASD (14 vs 7) patients accumulate several NS rare variants (Supplementary Figure 4).

Most promising NS rare variants

To identify the variants that are more likely to be disease-predisposing or -causing mutations, all the rare NS variants were ranked according to their potentially damaging effect on protein function and the mode of transmission and/or the cosegregation between the variants and the disease (Supplementary Table 7). The most promising variants are listed in Table 3 and several examples are described here.

We identified one potentially damaging variant c.733T>C (M: p.F245L) in *PPP1R3F*, which encodes a regulatory subunit of protein phosphatase 1, in one boy with an Asperger Syndrome diagnosis and seizures (Supplementary Figure 3a). This variant was transmitted from a mother who suffered from learning disabilities and seizures. Protein phosphatase 1 is a serine/threonine phosphatase that consists of a catalytic subunit that can interact with over 50 regulatory or scaffolding proteins and is involved in many cellular processes. In neurons, protein phosphatase 1 influences neurite outgrowth, synapse formation, and is critical for synaptic plasticity and learning and memory.⁶²

We also identified a missense variant, c.385G>A (M: p.G129R), in the 26S proteasome non-ATPase regulatory subunit (*PSMD10*) gene in one boy with classical autism. This variant was inherited from his heterozygous mother (Supplementary Figure 3b). The Gly129 residue is perfectly conserved among species and is located in the ANK3 (ankyrin) domain of PSMD10 (which is also known as gankyrin). Numerous synaptic proteins, including scaffolding proteins or neurotransmitter receptors, are regulated by degradation mechanisms through the ubiquitin–proteasome pathway, a phenomenon that is particularly important for synaptic plasticity.⁶³ Moreover, several proteins of this pathway, such as the ubiquitin ligase UBE3A, have been linked to autism-like disorders and MR.^{64,65} There is no direct evidence that PSMD10 is involved in the degradation of synaptic proteins; however, its synaptic localization is suggestive of this function. A non-conservative substitution of an amino acid in one ANK domain of this protein could alter its binding to protein partners and compromise its function.

The c.1845C>A (F + M: p.F555L) missense variant was identified in the gene encoding the MCF2/DBL protein, which is a member of the large family of GDP–GTP exchange factors for small GTPases of the Rho family (Supplementary Figure 3c). The F555L variant was found in one heterozygous SCZaff girl, her youngest SCZaff brother, and her unaffected mother. This variant affects a highly conserved amino acid and is predicted to be damaging to protein function, according to different programs (Supplementary Table 4). The Phe555 residue is located in the DBL motif, which is critical for the regulation of Rho-

Table 3 List of the 30 most promising NS rare variants identified during the screening of the 111 X-linked synaptic genes in 285 ASD and SCZ patients

Variant Info				Segregation analysis		Effect on protein function ^a
Gene	Variant			Gender	Pedigree ^a	
Rare missense variants in ASD patients						
PPP1R3F	T/C	Hemi	F245L	M	Mother + +	+ + +
MECP2	C/T	Hetero	R309W	F	<i>de novo</i> + +	+ +
PSMD10	G/A	Hemi	G129R	M	Mother +	+ +
PDZD4	G/A	Hemi	D326N	M	Mother +	+ +
GLRA2	G/T	Hetero	R350L	F	Mother	+ +
TM4SF2/TSPAN7	C/A	Hemi	P172H	M	Mother +	+
OPHN1	A/G	Hemi	H705R	M	Mother +	+
CXCR3	C/G	Hemi	V115L	M	Mother +	+
FRMPD4	G/A	Hemi	G1147E	M	Mother +	+
MAOA	A/T	Hemi	N271I	M	Mother +	+
PCDH19	C/G	Hemi	H146Q	M	Mother +	+
WNK3	C/G	Hemi	S886C	M	Mother +	+
WNK3	G/T	Hemi	G944V	M	Mother +	+
SH3KBP1	A/G	Hemi	S458G	M	Mother +	+
HNRPH2	G/C	Hemi	M396I	M	Mother +	+
Rare missense variants in SCZ patients						
MCF2	C/A	Hetero	F555L	F	Mother + +	+ +
GPRASP2	T/A	Hemi	I791K	M	Mother + +	+ +
FRMPD4	C/T	Hemi	R1189C	M	n.a. +	+ + +
HCFC1	C/T	Hetero	T142M	F	Mother + +	+ +
GABRE	C/T	Hetero	R472H	F	Mother	+ + +
OPHN1	A/G	Hemi	M461V	M	Mother +	+ +
BRS3	C/T	Hemi	R352W	M	Mother +	+ +
ATP2B3	G/C	Hemi	D459H	M	Mother +	+ +
BEX2	G/A	Hemi	R53H	M	Mother +	+ +
HTR2C	C/T	Hemi	A171V	M	Mother +	+ +
SLITRK2	G/A	Hetero	V89M	F	Mother + +	+
IDH3G	C/T	Hemi	R101C	M	Mother + +	+
PCDH19	T/C	Hetero	V91A	F	Mother	+ +
P2RY4	C/T	Hetero	S130N	F	Mother	+ +
SLC9A6	G/A	Hetero	R536Q	F	Mother	+ +

Abbreviations: ASD, autism spectrum disorder; F, female; Hemi, hemizygous; Hetero, heterozygous; M, male; n.a., non available; SCZ, schizophrenia.

^aMore details on pedigree (cosegregation analysis) or functional effect predictions are given in Supplementary Table 7.

GTPases. Other rare variants or polymorphisms were observed in this gene, but did not involve highly conserved residues. *Dbl* knockout mice reportedly display shorter dendrites in distinct populations of *Dbl*-null cortical pyramidal neurons, which suggests a function for MCF2/*Dbl* in dendrite elongation.⁶⁶ Reduced dendrite spines were observed in the prefrontal cortex of schizophrenic individuals, suggesting a possible involvement of the RhoGTPase pathway.^{18,19}

Other promising variants were identified in genes encoding proteins involved in cell adhesion (SLITRK2, PCDH11X, PCDH19, TM4SF2/TSPAN7), serotonin pathway (MAOA, HTR2C), mRNA or protein transport (HNRPH2, GPRASP2), calcium homeostasis (ATP2B3), or RhoGTPase regulation (OPHN1). Variants in several neurotransmitter receptors (GLRA2, GABRE, HTR2C) were also identified.

Finally, we identified variants in genes encoding proteins involved in various cell functions less specific of the synapse (WNK3, CXCR3, IDH3G, HCFC1 or SLC9A6) as well as proteins highly expressed in brain with unknown or unclear function (PDZD4, FRMPD4, BEX2).

Screening of MR genes

Sequence analysis of the *TM4SF2/TSPAN7* gene led to the identification of the P172H missense variant, which was associated earlier with MR in two families.^{67,68} In the first study, the cosegregation of this variant was not perfect, with one non-affected male carrying the variant. This variant was also identified by another team who screened patients with MR and was absent in 77 ASD patients⁶⁸ and 420 control chromosomes tested. We found the P172H variant in one ASD boy with MR; this variant was

inherited from his mother (Supplementary Figure 3d). We sequenced an additional 285 ASD patients for the fragment encoding this amino acid and found the same variant in another boy diagnosed with pervasive developmental disorder not otherwise specified without MR (Supplementary Table 5). This variant was neither present in 143 SCZ nor in 190 ethnically matched control individuals. Taken together, ours and earlier reports revealed that the P172H variant has not been found in 697 control X chromosomes. Our results add evidence to the function of the P172H substitution in *TM4SF2/TSPAN7* not only in MR but also in ASD. *TM4SF2/TSPAN7* encodes a cell-surface glycoprotein from the tetraspanin family, a class of proteins that have a function in the regulation of cell development, activation, growth, and motility. The Pro172 residue is located in the extracellular loop, a region that may be involved in interaction with other adhesion proteins.

The screening of the RTT *MECP2* gene revealed a *de novo* missense variant (F: p.R309W) in a girl with ASD. She presents with an important psychomotor retardation and had developed severe epileptic seizures from her adolescence. However, she does not satisfy the criteria for classical nor atypical RTT. This variant, already described in one boy with NS-MR,⁶⁹ affects a conserved amino acid located in a domain responsible for the interaction with the cyclin-dependent kinase-like CDKL5 and is predicted to be damaging.

We also identified two rare missense variants in *OPHN1*: c.1381A>G (M: p.M461V) in one COS patient and c.2114A>G (M: p.H705R) in one ASD patient (Supplementary Figure 3e). These two variants were predicted to affect protein function, according to different programs, and they target amino acids that are well conserved among species. In contrast, *OPHN1* polymorphisms and variants found in the control cohort (Supplementary Table 5) were all predicted to have a benign effect on protein function. *OPHN1* encodes a RhoGTPase-activating protein that interacts with RhoA, Rac, cdc42, actin filaments, and with the HOMER protein. It is expressed in both neuronal and glial cells and participates in the regulation of dendrite elongation.⁷⁰ Nonsense, splice, frameshift, and missense mutations have been described for this gene in patients with MR and cerebellar hypoplasia.⁷¹ Interestingly, cerebellar abnormalities have also been described in COS and ASD patients.^{72,73}

Among the other genes that are known to be involved in NS-MR, ASD, or ASD-related diseases, a subset showed no variation in their coding sequences, including *AP1S2*, *ARHGEF9*, *DCX*, *GDI1*, and *SLC6A8*. No NS variants were observed in *CASK* and *SYP*, two genes reported recently by Tarpey *et al.*⁵⁰ to cause NS-MR. In addition, no novel NS variants were identified during the screening of *ARHGEF6*, *ATP6AP2*, *ARX*, *JARID1C*, and *FMR1*. This was also the case for *RPL10*, in which variants were described recently in ASD patients.⁷⁴ Several

novel NS variants were identified in *CDKL5*, *DLG3*, *NLGN3*, *PAK3*, *AFF2*, and *RPS6KA3*; however, the analysis of the pedigrees and of the predicted effect of these variants on protein function suggests that none of them are likely to be involved in ASD.

Discussion

Our study reports the systematic resequencing of the coding regions of 111 X-linked candidate genes in a cohort of 285 ASD and SCZ patients, aimed at the identification of genes potentially involved in these disorders. Using various methods, we prioritized these genes from a larger group of 1125 X-linked genes and ~5000 potentially synaptic genes. For the X chromosome, only 28 genes were shared between the gene list from SynDB and that derived from proteomic studies. This non-concordance may be explained by three major reasons: (1) some synaptic proteins, such as membrane receptors, are difficult to purify and could be missed by proteomic studies, (2) the presence of a higher number of false positives in SynDB because of inclusion of genes based on structural predictions, and (3) a fraction of the proteins found through proteomic studies may also represent false positives, as they could be due to contamination of the synaptic fraction during the preparation. These results convinced us to use the complementarity of these two types of sources. Altogether, the gene lists extracted from SynDB and from proteomic studies comprised 65% of the genes identified through our manual search of PubMed. The combination of these three resources resulted in a comprehensive list of X chromosome potentially synaptic genes. This list also provides good candidate genes for the screening of other psychiatric and neurological disorders, such as bipolar disorder or Tourette's syndrome.

Here, we identified several hundred variants, the majority of which had not been described in any public database. Tarpey *et al.*,⁵⁰ who recently published a list of X-chromosome variants found in patients with MR, identified 134 of the variants reported here, which corresponded mainly to the polymorphism category. However, their cohort is not a pertinent control for our study, as MR and ASD are genetically related and are often clinically comorbid. We screened three times less nucleotides (75 954 657 bp vs ~200 000 000 bp) than Tarpey *et al.* and found around three times less variants in coding regions (533 vs 1858). However, we identified significantly less truncating variants (6 vs 40). This result is not surprising, as their cohort included only families with X-linked MR. Consistent with their results, we observed that truncating variants in two genes (*P2RY4* and *HS6ST2*) seemed to be well tolerated in males and to not lead to any obvious phenotype. Contrary to the variant in *P2RY4*, which removes only the C-terminal end of the protein, the variant in *HS6ST2* removes almost all the protein. Possible functional redundancy with other heparan

sulfate 6-O-sulfotransferase subtypes, such as HS6ST1 and HS6ST3, may explain the fact that HS6ST2 truncation is tolerated in male individuals.

Despite the probable function of rare variants with a recent origin in ASD and SCZ, we identified and confirmed only two *de novo* damaging variants, which were found in *IL1RAPL1* and *MECP2* in two ASD girls. However, because of the particularity of X-linked transmission, we are aware that X-linked genes are not ideal to test the *de novo* hypothesis. One nonsense and several missense variants were identified in the *MAOB* gene in SCZ patients. *MAOB* catalyzes the oxidative deamination of xenobiotic and biogenic amines, such as the neurotransmitter dopamine or the neuromodulator phenylethylamine.⁷⁵ Phenylethylamine is involved in the modulation of mood and, as it is structurally close to amphetamine, it may cause, when expressed at high levels, a similar type of psychosis.⁷⁶ Moreover, low *MAOB* activity and elevated phenylethylamine in urine were described earlier in SCZ patients.⁷⁷ Several studies failed to identify any association between polymorphisms in *MAOB* and SCZ; however, one positive association was recently found in the Spanish population.⁷⁸ Nonetheless, one would not expect positive associations if disease-predisposing mutations are different rare mutations. Although no evident neurological phenotype was found in two boys who carry a deletion of a fragment of *MAOB* and of the Norrie disease gene,⁷⁹ *MAOB* knockout mice present elevated phenylethylamine in urine and an increased reactivity to stress.⁸⁰ Regarding the possible implication of *MAOB* in SCZ, the identification of one nonsense mutation in this gene in one SCZ patient is potentially relevant, even if the cosegregation was not perfect. Further studies are needed to determine the effect of the three missense variants found in the four additional families on *MAOB* function.

The NS rare variants identified in this study were predicted to be more damaging than the NS polymorphisms, which suggests that they may include disease-causing mutations responsible for ASD or SCZ. The genetic follow-up on these variants is challenging, as most of our patients are sporadic cases. Moreover, because of the genetic complexity of ASD and SCZ, we do not expect to find a perfect segregation of variants with the disease, which hampers the interpretation of the results. Several genes in which we identified potentially damaging variants have a function in structural modulation of the synapse. The involvement of the *IL1RAPL1*, *OPHN1*, and *MCF2* proteins in neurite outgrowth is well documented^{66,70,81} and evidence suggests that *SLITRK2* and *TM4SF2/TSPAN7* may also be implicated in this process.^{56,67} We also highlighted variants in genes involved in serotonin function (*HTR2C*) and degradation (*MAOB*) in SCZ patients.

Our gene scoring system aimed at selecting the best candidate genes seems to have performed well. *MAOB* was selected with a score equal to 6, which was the maximum score. Several of the genes in

which we found potentially relevant variants in ASD (for example *IL1RAPL1*, *OPHN1*, *TM4SF2/TSPAN7*, or *MECP2*) were selected because of their involvement in NS-MR. Interestingly, most of our ASD individuals who carried mutations in these genes did not have MR. These results support a genetic link between MR and ASD, suggesting that the same genes and mutations may predispose to these two diseases, as was shown earlier for *NLGN4*.¹² Therefore, the screening of genes involved in MR in ASD patients is warranted, even if the patients do not have MR. The study of factors (environmental or genetic) that modulate these phenotypes will be critical to understand the molecular pathways that underlie MR, ASD, or both.

The identification of two damaging variants in *OPHN1* in one male with ASD and in another with COS suggests a genetic link between these two diseases. The COS patient was also diagnosed with pervasive developmental disorder as one-third of the individuals with COS. The hypothesis of the existence of common genes between ASD and SCZ is emerging⁸² and has been shown by the discovery that copy number variants of the *NRXN1* synaptic gene are associated with these two diseases.⁸³ Similarly, the SCZ-associated *DISC-1* gene was also found to be involved in ASD.⁸⁴ Interestingly, in concordance with our results, a recent study showed that several copy number variants are common between MR, ASD, and SCZ, supporting the existence of shared biologic pathways in these neurodevelopmental disorders.^{85,86}

In conclusion, our results indicate that large-scale direct resequencing of synaptic candidate genes constitutes a promising approach to dissect the genetic heterogeneity of SCZ and ASD and to explore the hypothesis that a number of distinct individually rare penetrant variants are involved in the pathogenesis of these two syndromes. Indeed, the identification of an excess of potentially damaging rare variants in ASD and SCZ patients validated the usefulness of this approach. However, with the exception of mutations in *IL1RAPL1*, *MECP2* and maybe *TM4SF2/TSPAN7* in ASD, it is difficult to make a definitive claim that the damaging variants identified here are disease causing. The truncating mutation in *MAOB*, as well as other missense variants in different genes are promising, but further work involving functional testing will be needed to confirm the implication of these rare variants in SCZ and ASD. If the hypothesis of rare variants with classical Mendelian inheritance does not seem to entirely explain the complexity of the genetic factors involved in ASD, we succeeded, at least, in the identification of causative mutations in some ASD patients. However, the story seems to be more complicated for SCZ. The excess of patients accumulating several NS rare variants in the SCZ cohort in comparison with the ASD cohort and the lack of identification of a clear segregating damaging mutation in SCZ in our study suggest that SCZ may involve more an interaction between rare variants with moderate effects in

different synaptic genes rather than addition of Mendelian inheritances. Interestingly, a recent genome-wide association study on SCZ reported that common polygenic variants could also contribute to the risk of SCZ.⁷ These common variants may work in concert with rare variants to manifest SCZ. However, we are limited in this conclusion because we analyzed only variants on the X chromosome. A similar approach targeting autosomal genes is needed to examine whether autosomal rare variants provide a similar picture to that of our X-chromosome study. More generally, for these diseases, the 'rare variant with penetrant effect' vs the 'common variant with low effect' hypotheses should not be viewed as exclusive hypotheses, but more as a continuum including also variants with rare frequency and having moderate effect. That is why direct resequencing of candidate genes, as well as copy number variants or genome-wide association study analyses, could be viewed as complementary approaches to dissect the genetic susceptibilities to SCZ and ASD.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

We thank the families involved in our study and the recruitment coordinators (Anne Desjarlais, Caroline Poulin, and Sabrina Diab). We thank Annie Levert, Judith St-Onge, and Isabelle Bachand for performing DNA extraction and paternity and identity testing. We are thankful for the efforts of the members of the McGill University and Genome Quebec Innovation Centre Sequencing (Pierre Lepage, Sébastien Brunet, and Hao Fan Yam) and Bioinformatic (Louis Létourneau and Louis Dumond Joseph) groups. This work was supported by a grant from Genome Canada and Génome Québec and was cofunded by Université de Montréal, for the 'Synapse-to-disease' (S2D) project.

References

- Cardno AG, Gottesman II. Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics. *Am J Med Genet* 2000; **97**: 12–17.
- Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E *et al*. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* 1995; **25**: 63–77.
- Sullivan PF. The genetics of schizophrenia. *PLoS Med* 2005; **2**: e212.
- Harrison PJ, Weinberger DR. Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. *Mol Psychiatry* 2005; **10**: 40–68; image 45.
- Yang MS, Gill M. A review of gene linkage, association and expression studies in autism and an assessment of convergent evidence. *Int J Dev Neurosci* 2007; **25**: 69–85.
- Ma D, Salyakina D, Jaworski JM, Konidari I, Whitehead PL, Andersen AN *et al*. A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Ann Hum Genet* 2009; **73**(Part 3): 263–273.
- Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF *et al*. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
- Gauthier J, Spiegelman D, Piton A, Lafreniere RG, Laurent S, St-Onge J *et al*. Novel *de novo* SHANK3 mutation in autistic patients. *Am J Med Genet B Neuropsychiatr Genet* 2009; **150B**: 421–424.
- Jamain S, Quach H, Betancur C, Rastam M, Colineaux C, Gillberg IC *et al*. Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet* 2003; **34**: 27–29.
- Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F *et al*. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet* 2007; **39**: 25–27.
- Moessner R, Marshall CR, Sutcliffe JS, Skaug J, Pinto D, Vincent J *et al*. Contribution of SHANK3 mutations to autism spectrum disorder. *Am J Hum Genet* 2007; **81**: 1289–1297.
- Laumonnier F, Bonnet-Brilhaut F, Gomot M, Blanc R, David A, Moizard MP *et al*. X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am J Hum Genet* 2004; **74**: 552–557.
- Zoghbi HY. Postnatal neurodevelopmental disorders: meeting at the synapse? *Science* 2003; **302**: 826–830.
- Garber K. Neuroscience. Autism's cause may reside in abnormalities at the synapse. *Science* 2007; **317**: 190–191.
- Harrison PJ. The hippocampus in schizophrenia: a review of the neuropathological evidence and its pathophysiological implications. *Psychopharmacology (Berl)* 2004; **174**: 151–162.
- Hashimoto R, Tankou S, Takeda M, Sawa A. Postsynaptic density: a key convergent site for schizophrenia susceptibility factors and possible target for drug development. *Drugs Today (Barc)* 2007; **43**: 645–654.
- Eastwood SL. The synaptic pathology of schizophrenia: is aberrant neurodevelopment and plasticity to blame? *Int Rev Neurobiol* 2004; **59**: 47–72.
- Garey LJ, Ong WY, Patel TS, Kanani M, Davis A, Mortimer AM *et al*. Reduced dendritic spine density on cerebral cortical pyramidal neurons in schizophrenia. *J Neurol Neurosurg Psychiatry* 1998; **65**: 446–453.
- Hill JJ, Hashimoto T, Lewis DA. Molecular mechanisms contributing to dendritic spine alterations in the prefrontal cortex of subjects with schizophrenia. *Mol Psychiatry* 2006; **11**: 557–566.
- Marco EJ, Skuse DH. Autism-lessons from the X chromosome. *Soc Cogn Affect Neurosci* 2006; **1**: 183–193.
- Loat CS, Haworth CM, Plomin R, Craig IW. A model incorporating potential skewed X-inactivation in MZ girls suggests that X-linked QTLs exist for several social behaviours including autism spectrum disorder. *Ann Hum Genet* 2008; **72**(Part 6): 742–751.
- Laumonnier F, Cuthbert PC, Grant SG. The role of neuronal complexes in human X-linked brain diseases. *Am J Hum Genet* 2007; **80**: 205–220.
- Skuse DH. X-linked genes and mental functioning. *Hum Mol Genet* 2005; **4 Spec No. 1**: R27–R32.
- Skuse D. Genetic influences on the neural basis of social cognition. *Philos Trans R Soc Lond B Biol Sci* 2006; **361**: 2129–2141.
- Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A *et al*. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 1991; **65**: 905–914.
- Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999; **23**: 185–188.
- Yonan AL, Alarcon M, Cheng R, Magnusson PK, Spence SJ, Palmer AA *et al*. A genomewide screen of 345 families for autism-susceptibility loci. *Am J Hum Genet* 2003; **73**: 886–897.
- Petit E, Herault J, Raynaud M, Cherpi C, Perrot A, Barthelemy C *et al*. X chromosome and infantile autism. *Biol Psychiatry* 1996; **40**: 457–464.
- Vincent JB, Melmer G, Bolton PF, Hodgkinson S, Holmes D, Curtis D *et al*. Genetic linkage analysis of the X chromosome in autism, with emphasis on the fragile X region. *Psychiatr Genet* 2005; **15**: 83–90.

- 30 Liu J, Nyholt DR, Magnussen P, Parano E, Pavone P, Geschwind D et al. A genomewide screen for autism susceptibility loci. *Am J Hum Genet* 2001; **69**: 327–340.
- 31 Auranen M, Vanhala R, Varilo T, Ayers K, Kempas E, Ylisaukko-Oja T et al. A genomewide screen for autism-spectrum disorders: evidence for a major susceptibility locus on chromosome 3q25-27. *Am J Hum Genet* 2002; **71**: 777–790.
- 32 Gauthier J, Joobor R, Dube MP, St-Onge J, Bonnel A, Garipey D et al. Autism spectrum disorders associated with X chromosome markers in French-Canadian males. *Mol Psychiatry* 2006; **11**: 206–213.
- 33 Dann J, DeLisi LE, Devoto M, Laval S, Nancarrow DJ, Shields G et al. A linkage study of schizophrenia to markers within Xp11 near the MAOB gene. *Psychiatry Res* 1997; **70**: 131–143.
- 34 DeLisi LE, Wellman N, Stewart J, Smith AB, Churchman M, Crow TJ. Linkage disequilibrium study of markers within the pericentromeric region of the X chromosome. *Am J Med Genet* 1999; **88**: 588–589.
- 35 Hovatta I, Varilo T, Suvisaari J, Terwilliger JD, Ollikainen V, Arajärvi R et al. A genomewide screen for schizophrenia genes in an isolated Finnish subpopulation, suggesting multiple susceptibility loci. *Am J Hum Genet* 1999; **65**: 1114–1124.
- 36 Wei J, Hemmings GP. A further study of a possible locus for schizophrenia on the X chromosome. *Biochem Biophys Res Commun* 2006; **344**: 1241–1245.
- 37 Roser P, Kawohl W. Turner syndrome and schizophrenia: a further hint for the role of the X-chromosome in the pathogenesis of schizophrenic disorders. *World J Biol Psychiatry* 2008; Sept 11: 1–4.
- 38 DeLisi LE, Friedrich U, Wahlstrom J, Boccio-Smith A, Forsman A, Eklund K et al. Schizophrenia and sex chromosome anomalies. *Schizophr Bull* 1994; **20**: 495–505.
- 39 El Abd S, Patton MA, Turk J, Hoey H, Howlin P. Social, communicational, and behavioral deficits associated with ring X turner syndrome. *Am J Med Genet* 1999; **88**: 510–516.
- 40 Fombonne E. Epidemiological trends in rates of autism. *Mol Psychiatry* 2002; **7** Suppl 2: S4–S6.
- 41 Goldstein JM, Seidman LJ, Goodman JM, Koren D, Lee H, Weintraub S et al. Are there sex differences in neuropsychological functions among patients with schizophrenia? *Am J Psychiatry* 1998; **155**: 1358–1364.
- 42 Trinidad JC, Specht CG, Thalhammer A, Schoepfer R, Burlingame AL. Comprehensive identification of phosphorylation sites in postsynaptic density preparations. *Mol Cell Proteomics* 2006; **5**: 914–922.
- 43 Takamori S, Holt M, Stenius K, Lemke EA, Grønborg M, Riedel D et al. Molecular anatomy of a trafficking organelle. *Cell* 2006; **127**: 831–846.
- 44 Collins MO, Husi H, Yu L, Brandon JM, Anderson CN, Blackstock WP et al. Molecular characterization and comparison of the components and multiprotein complexes in the postsynaptic proteome. *J Neurochem* 2006; April; **97** Suppl 1: 16–23.
- 45 Zhang W, Zhang Y, Zheng H, Zhang C, Xiong W, Olyarchuk JG et al. SynDB: a Synapse protein DataBase based on synapse ontology. *Nucleic Acids Res* 2007; **35**(Database issue): D737–D741.
- 46 Gauthier J, Bonnel A, St-Onge J, Karemera L, Laurent S, Mottron L et al. NLGN3/NLGN4 gene mutations are not responsible for autism in the Quebec population. *Am J Med Genet B Neuropsychiatr Genet* 2005; **132B**: 74–75.
- 47 Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N et al. PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res* 2003; **31**: 334–341.
- 48 Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res* 2002; **30**: 3894–3900.
- 49 Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003; **31**: 3812–3814.
- 50 Tarpey PS, Smith R, Pleasance E, Whibley A, Edkins S, Hardy C et al. A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat Genet* 2009; **41**: 535–543.
- 51 Piton A, Michaud JL, Peng H, Aradhya S, Gauthier J, Mottron L et al. Mutations in the calcium-related gene IL1RAPL1 are associated with autism. *Hum Mol Genet* 2008; **17**: 3965–3974.
- 52 Binda C, Newton-Vinson P, Hubalek F, Edmondson DE, Mattevi A. Structure of human monoamine oxidase B, a drug target for the treatment of neurological disorders. *Nat Struct Biol* 2002; **9**: 22–26.
- 53 Ng PC, Levy S, Huang J, Stockwell TB, Walenz BP, Li K et al. Genetic variation in an individual human exome. *PLoS Genet* 2008; **4**: e1000160.
- 54 Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 2002; **3**: 285–298.
- 55 Collin RW, de Heer AM, Oostrik J, Pauw RJ, Plantinga RF, Huygen PL et al. Mid-frequency DFNA8/12 hearing loss caused by a synonymous TECTA mutation that affects an exonic splice enhancer. *Eur J Hum Genet* 2008; **16**: 1430–1436.
- 56 Aruga J, Mikoshiba K. Identification and characterization of Slitrk, a novel neuronal transmembrane protein family controlling neurite outgrowth. *Mol Cell Neurosci* 2003; **24**: 117–129.
- 57 Hemara-Wahani A, Berjukow S, Hope CI, Dearden PK, Wu SB, Wilson-Wheeler J et al. A CACNA1F mutation identified in an X-linked retinal disorder shows the voltage dependence of Cav1.4 channel activation. *Proc Natl Acad Sci USA* 2005; **102**: 7553–7558.
- 58 Splawski I, Timothy KW, Sharpe LM, Decher N, Kumar P, Bloise R et al. Ca(V)1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell* 2004; **119**: 19–31.
- 59 Splawski I, Yoo DS, Stotz SC, Cherry A, Clapham DE, Keating MT. CACNA1H mutations in autism spectrum disorders. *J Biol Chem* 2006; **281**: 22085–22091.
- 60 Strom SP, Stone JL, Ten Bosch JR, Merriman B, Cantor RM, Geschwind DH et al. High-density SNP association study of the 17q21 chromosomal region linked to autism identifies CACNA1G as a novel candidate gene. *Mol Psychiatry* 19 May 2009; e-pub ahead of print.
- 61 Weaving LS, Christodoulou J, Williamson SL, Friend KL, McKenzie OL, Archer H et al. Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *Am J Hum Genet* 2004; **75**: 1079–1093.
- 62 Munton RP, Vizi S, Mansuy IM. The role of protein phosphatase-1 in the modulation of synaptic and structural plasticity. *FEBS Lett* 2004; **567**: 121–128.
- 63 Bingol B, Schuman EM. Synaptic protein degradation by the ubiquitin proteasome system. *Curr Opin Neurobiol* 2005; **15**: 536–541.
- 64 Kishino T, Lalande M, Wagstaff J. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat Genet* 1997; **15**: 70–73.
- 65 Nurmi EL, Bradford Y, Chen Y, Hall J, Arnone B, Gardiner MB et al. Linkage disequilibrium at the Angelman syndrome gene UBE3A in autism families. *Genomics* 2001; **77**: 105–113.
- 66 Hirsch E, Pozzato M, Vercelli A, Barberis L, Azzolino O, Russo C et al. Defective dendrite elongation but normal fertility in mice lacking the Rho-like GTPase activator Dbl. *Mol Cell Biol* 2002; **22**: 3140–3148.
- 67 Zemni R, Bienvenu T, Vinet MC, Sefiani A, Carrie A, Billuart P et al. A new gene involved in X-linked mental retardation identified by analysis of an X;2 balanced translocation. *Nat Genet* 2000; **24**: 167–170.
- 68 Maranduba CM, Sa Moreira E, Muller Orabona G, Pavanetto RC, Vianna-Morgante AM, Passos-Bueno MR. Does the P172H mutation at the TM4SF2 gene cause X-linked mental retardation? *Am J Med Genet A* 2004; **124A**: 413–415.
- 69 Campos Jr M, Abdalla CB, Santos-Reboucas CB, dos Santos AV, Pestana CP, Domingues ML et al. Low significance of MECP2 mutations as a cause of mental retardation in Brazilian males. *Brain Dev* 2007; **29**: 293–297.
- 70 Govek EE, Newey SE, Akerman CJ, Cross JR, Van der Veken L, Van Aelst L. The X-linked mental retardation protein oligophrenin-1 is required for dendritic spine morphogenesis. *Nat Neurosci* 2004; **7**: 364–372.
- 71 Zanni G, Saillour Y, Nagara M, Billuart P, Castelnau L, Moraine C et al. Oligophrenin 1 mutations frequently cause X-linked mental retardation with cerebellar hypoplasia. *Neurology* 2005; **65**: 1364–1369.
- 72 Mittleman G, Goldowitz D, Heck DH, Blaha CD. Cerebellar modulation of frontal cortex dopamine efflux in mice: relevance to autism and schizophrenia. *Synapse* 2008; **62**: 544–550.

- 73 Kaufmann WE, Cooper KL, Mostofsky SH, Capone GT, Kates WR, Newschaffer CJ *et al*. Specificity of cerebellar vermal abnormalities in autism: a quantitative magnetic resonance imaging study. *J Child Neurol* 2003; **18**: 463–470.
- 74 Klauck SM, Felder B, Kolb-Kokocinski A, Schuster C, Chiocchetti A, Schupp I *et al*. Mutations in the ribosomal protein gene RPL10 suggest a novel modulating disease mechanism for autism. *Mol Psychiatry* 2006; **11**: 1073–1084.
- 75 Shih JC, Chen K, Ridd MJ. Role of MAO A and B in neurotransmitter metabolism and behavior. *Pol J Pharmacol* 1999; **51**: 25–29.
- 76 O'Reilly RL, Davis BA. Phenylethylamine and schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 1994; **18**: 63–75.
- 77 Wyatt RJ, Potkin SG, Bridge TP, Phelps BH, Wise CD. Monoamine oxidase in schizophrenia: an overview. *Schizophr Bull* 1980; **6**: 199–207.
- 78 Gasso P, Bernardo M, Mas S, Crescenti A, Garcia C, Parellada E *et al*. Association of A/G polymorphism in intron 13 of the monoamine oxidase B gene with schizophrenia in a Spanish population. *Neuropsychobiology* 2008; **58**: 65–70.
- 79 Lenders JW, Eisenhofer G, Abeling NG, Berger W, Murphy DL, Konings CH *et al*. Specific genetic deficiencies of the A and B isoenzymes of monoamine oxidase are characterized by distinct neurochemical and clinical phenotypes. *J Clin Invest* 1996; **97**: 1010–1019.
- 80 Grimsby J, Toth M, Chen K, Kumazawa T, Klaidman L, Adams JD *et al*. Increased stress response and beta-phenylethylamine in MAOB-deficient mice. *Nat Genet* 1997; **17**: 206–210.
- 81 Gambino F, Pavlowsky A, Begle A, Dupont JL, Bahi N, Courjaret R *et al*. IL1-receptor accessory protein-like 1 (IL1RAPL1), a protein involved in cognitive functions, regulates N-type Ca²⁺ -channel and neurite elongation. *Proc Natl Acad Sci USA* 2007; **104**: 9063–9068.
- 82 Burbach JP, van der Zwaag B. Contact in the genetics of autism and schizophrenia. *Trends Neurosci* 2009; **32**: 69–72.
- 83 Kirov G, Gumus D, Chen W, Norton N, Georgieva L, Sari M *et al*. Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet* 2008; **17**: 458–465.
- 84 Kilpinen H, Ylisaukko-Oja T, Hennah W, Palo OM, Varilo T, Vanhala R *et al*. Association of DISC1 with autism and Asperger syndrome. *Mol Psychiatry* 2008; **13**: 187–196.
- 85 Guilmatre A, Dubourg C, Mosca AL, Legallie S, Goldenberg A, Drouin-Garraud V *et al*. Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. *Arch Gen Psychiatry* 2009; **66**: 947–956.
- 86 Sebat J, Levy DL, McCarthy SE. Rare structural variants in schizophrenia: one disorder, multiple mutations; one mutation, multiple disorders. *Trends Genet* 2009; **25**: 528–535.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (<http://www.nature.com/mp>)

Bibliographie

2008. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 455:237-41
- Akbarian S, Sucher NJ, Bradley D, Tafazzoli A, Trinh D, et al. 1996. Selective alterations in gene expression for NMDA receptor subunits in prefrontal cortex of schizophrenics. *J Neurosci* 16:19-30
- Allen AJ, Griss ME, Folley BS, Hawkins KA, Pearlson GD. 2009. Endophenotypes in schizophrenia: a selective review. *Schizophr Res* 109:24-37
- Allen NC, Bagade S, McQueen MB, Ioannidis JP, Kavvoura FK, et al. 2008. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* 40:827-34
- Amitai N, Markou A. 2010. Disruption of performance in the five-choice serial reaction time task induced by administration of N-methyl-D-aspartate receptor antagonists: relevance to cognitive dysfunction in schizophrenia. *Biol Psychiatry* 68:5-16
- Athanasios L, Mattingsdal M, Kahler AK, Brown A, Gustafsson O, et al. 2010. Gene variants associated with schizophrenia in a Norwegian genome-wide study are replicated in a large European cohort. *J Psychiatr Res* 44:748-53
- Awadalla P, Gauthier J, Myers RA, Casals F, Hamdan FF, et al. 2010. Direct Measure of the De Novo Mutation Rate in Autism and Schizophrenia Cohorts. *Am J Hum Genet*
- Bading H, Ginty DD, Greenberg ME. 1993. Regulation of gene expression in hippocampal neurons by distinct calcium signaling pathways. *Science* 260:181-6
- Badner JA, Gershon ES. 2002. Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. *Mol Psychiatry* 7:405-11
- Banerjee A, Macdonald ML, Borgmann-Winter KE, Hahn CG. 2010. Neuregulin 1-erbB4 pathway in schizophrenia: From genes to an interactome. *Brain Res Bull*
- Bassett AS, Bury A, Hodgkinson KA, Honer WG. 1996. Reproductive fitness in familial schizophrenia. *Schizophr Res* 21:151-60
- Bassett AS, Scherer SW, Brzustowicz LM. 2010. Copy number variations in schizophrenia: critical review and new perspectives on concepts of genetics and disease. *Am J Psychiatry* 167:899-914
- Belforte JE, Zsiros V, Sklar ER, Jiang Z, Yu G, et al. 2010. Postnatal NMDA receptor ablation in corticolimbic interneurons confers schizophrenia-like phenotypes. *Nat Neurosci* 13:76-83
- Blackwood DH, Thiagarajah T, Malloy P, Pickard BS, Muir WJ. 2008. Chromosome abnormalities, mental retardation and the search for genes in bipolar disorder and schizophrenia. *Neurotox Res* 14:113-20
- Brown AS, Derkits EJ. 2010. Prenatal infection and schizophrenia: a review of epidemiologic and translational studies. *Am J Psychiatry* 167:261-80
- Byrne M, Agerbo E, Bennedsen B, Eaton WW, Mortensen PB. 2007. Obstetric conditions and risk of first admission with schizophrenia: a Danish national register based study. *Schizophr Res* 97:51-9
- Cannon M, Jones PB, Murray RM. 2002. Obstetric complications and schizophrenia: historical and meta-analytic review. *Am J Psychiatry* 159:1080-92
- Cheli V, Adrover M, Blanco C, Ferrari C, Cornea A, et al. 2006. Knocking-down the NMDAR1 subunit in a limited amount of neurons in the rat hippocampus impairs learning. *J Neurochem* 97 Suppl 1:68-73

- Chubb JE, Bradshaw NJ, Soares DC, Porteous DJ, Millar JK. 2008. The DISC locus in psychiatric illness. *Mol Psychiatry* 13:36-64
- Collingridge GL, Isaac JT, Wang YT. 2004. Receptor trafficking and synaptic plasticity. *Nat Rev Neurosci* 5:952-62
- Conrad DF, Keebler JE, Depristo MA, Lindsay SJ, Zhang Y, et al. 2011. Variation in genome-wide mutation rates within and between human families. *Nat Genet*
- Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, et al. 2010. Origins and functional impact of copy number variation in the human genome. *Nature* 464:704-12
- Covarrubias M, Bhattacharji A, De Santiago-Castillo JA, Dougherty K, Kaulin YA, et al. 2008. The neuronal Kv4 channel complex. *Neurochem Res* 33:1558-67
- Craddock N, Owen MJ. 2010. The Kraepelinian dichotomy - going, going... but still not gone. *Br J Psychiatry* 196:92-5
- Crow TJ. 2007. How and why genetic linkage has not solved the problem of psychosis: review and hypothesis. *Am J Psychiatry* 164:13-21
- Crow TJ. 2011. 'The missing genes: what happened to the heritability of psychiatric disorders?'. *Mol Psychiatry* 16:362-4
- Cull-Candy SG, Leszkiewicz DN. 2004. Role of distinct NMDA receptor subtypes at central synapses. *Sci STKE* 2004:re16
- Dalmau J, Gleichman AJ, Hughes EG, Rossi JE, Peng X, et al. 2008. Anti-NMDA-receptor encephalitis: case series and analysis of the effects of antibodies. *Lancet Neurol* 7:1091-8
- Dalmau J, Lancaster E, Martinez-Hernandez E, Rosenfeld MR, Balice-Gordon R. 2011. Clinical experience and laboratory investigations in patients with anti-NMDAR encephalitis. *Lancet Neurol* 10:63-74
- Das S, Sasaki YF, Rothe T, Premkumar LS, Takasu M, et al. 1998. Increased NMDA current and spine density in mice lacking the NMDA receptor subunit NR3A. *Nature* 393:377-81
- Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, et al. 2007. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet* 39:25-7
- Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, et al. 2010. A map of human genome variation from population-scale sequencing. *Nature* 467:1061-73
- Endele S, Rosenberger G, Geider K, Popp B, Tamer C, et al. 2010. Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat Genet* 42:1021-6
- Esslinger C, Walter H, Kirsch P, Erk S, Schnell K, et al. 2009. Neural mechanisms of a genome-wide supported psychosis variant. *Science* 324:605
- Ewald RC, Cline HT. 2009. NMDA Receptors and Brain Development.
- Gao XM, Sakai K, Roberts RC, Conley RR, Dean B, Tamminga CA. 2000. Ionotropic glutamate receptors and expression of N-methyl-D-aspartate receptor subunits in subregions of human hippocampus: effects of schizophrenia. *Am J Psychiatry* 157:1141-9
- Gauthier J, Champagne N, Lafreniere RG, Xiong L, Spiegelman D, et al. 2010. De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for schizophrenia. *Proc Natl Acad Sci U S A* 107:7863-8

- Gauthier J, Siddiqui TJ, Huashan P, Yokomaku D, Hamdan FF, et al. 2011. Truncating mutations in NRXN2 and NRXN1 in autism spectrum disorders and schizophrenia. *Hum Genet*
- Gauthier J, Spiegelman D, Piton A, Lafreniere RG, Laurent S, et al. 2009. Novel de novo SHANK3 mutation in autistic patients. *Am J Med Genet B Neuropsychiatr Genet* 150B:421-4
- Girard S, Gauthier J, Noreau N, et al. Increased exonic de novo mutation rate in probands affected with schizophrenia : Nature genetics, in press
- Glantz LA, Gilmore JH, Lieberman JA, Jarskog LF. 2006. Apoptotic mechanisms and the synaptic pathology of schizophrenia. *Schizophr Res* 81:47-63
- Glessner JT, Reilly MP, Kim CE, Takahashi N, Albano A, et al. 2010. Strong synaptic transmission impact by copy number variations in schizophrenia. *Proc Natl Acad Sci U S A* 107:10584-9
- Gottesman, II, Gould TD. 2003. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* 160:636-45
- Gray JA, Roth BL. 2007. The pipeline and future of drug development in schizophrenia. *Mol Psychiatry* 12:904-22
- Green T, Gothelf D, Glaser B, Debbane M, Frisch A, et al. 2009. Psychiatric disorders and intellectual functioning throughout development in velocardiofacial (22q11.2 deletion) syndrome. *J Am Acad Child Adolesc Psychiatry* 48:1060-8
- Guillaud L, Setou M, Hirokawa N. 2003. KIF17 dynamics and regulation of NR2B trafficking in hippocampal neurons. *J Neurosci* 23:131-40
- Guilmatre A, Dubourg C, Mosca AL, Legallic S, Goldenberg A, et al. 2009. Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. *Arch Gen Psychiatry* 66:947-56
- Hahn CG, Wang HY, Cho DS, Talbot K, Gur RE, et al. 2006. Altered neuregulin 1-erbB4 signaling contributes to NMDA receptor hypofunction in schizophrenia. *Nat Med* 12:824-8
- Hamdan FF, Daoud H, Rochefort D, Piton A, Gauthier J, et al. 2010. De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. *Am J Hum Genet* 87:671-8
- Hamdan FF, Gauthier J, Araki Y, Lin DT, Yoshizawa Y, et al. 2011. Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. *Am J Hum Genet* 88:306-16
- Hamdan FF, Gauthier J, Spiegelman D, Noreau A, Yang Y, et al. 2009. Mutations in SYNGAP1 in autosomal nonsyndromic mental retardation. *N Engl J Med* 360:599-605
- Harrison PJ, Weinberger DR. 2005. Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. *Mol Psychiatry* 10:40-68; image 5
- Haukka J, Suvisaari J, Lonnqvist J. 2003. Fertility of patients with schizophrenia, their siblings, and the general population: a cohort study from 1950 to 1959 in Finland. *Am J Psychiatry* 160:460-3
- Hayashi-Takagi A, Sawa A. 2010. Disturbed synaptic connectivity in schizophrenia: convergence of genetic risk factors during neurodevelopment. *Brain Res Bull* 83:140-6
- Higgins J. 1966. Effects of child rearing by schizophrenic mothers. *J Psychiatr Res* 4:153-67

- Higgins J. 1976. Effects of child rearing by schizophrenic mothers: a follow-up. *J Psychiatr Res* 13:1-9
- Hirokawa N, Niwa S, Tanaka Y. 2010. Molecular motors in neurons: transport mechanisms and roles in brain function, development, and disease. *Neuron* 68:610-38
- Hirokawa N, Takemura R. 2005. Molecular motors and mechanisms of directional transport in neurons. *Nat Rev Neurosci* 6:201-14
- Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, et al. 2004. Detection of large-scale variation in the human genome. *Nat Genet* 36:949-51
- Ibrahim HM, Hogg AJ, Jr., Healy DJ, Haroutunian V, Davis KL, Meador-Woodruff JH. 2000. Ionotropic glutamate receptor binding and subunit mRNA expression in thalamic nuclei in schizophrenia. *Am J Psychiatry* 157:1811-23
- Ikeda M, Aleksic B, Kinoshita Y, Okochi T, Kawashima K, et al. 2011. Genome-wide association study of schizophrenia in a Japanese population. *Biol Psychiatry* 69:472-8
- Inada K, Ishigooka J, Anzai T, Suzuki E, Miyaoka H, Saji M. 2003. Antisense hippocampal knockdown of NMDA-NR1 by HVJ-liposome vector induces deficit of prepulse inhibition but not of spatial memory. *Neurosci Res* 45:473-81
- Jacquet H, Raux G, Thibaut F, Hecketsweiler B, Houy E, et al. 2002. PRODH mutations and hyperprolinemia in a subset of schizophrenic patients. *Hum Mol Genet* 11:2243-9
- Jarskog LF, Glantz LA, Gilmore JH, Lieberman JA. 2005. Apoptotic mechanisms in the pathophysiology of schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 29:846-58
- Javitt DC, Zukin SR. 1991. Recent advances in the phencyclidine model of schizophrenia. *Am J Psychiatry* 148:1301-8
- Jia P, Sun J, Guo AY, Zhao Z. 2010. SZGR: a comprehensive schizophrenia gene resource. *Mol Psychiatry* 15:453-62
- Johansson AC, Feuk L. 2011. Characterization of copy number stable regions in the human genome. *Hum Mutat*
- Jolin EM, Weller RA, Weller EB. 2011. Occurrence of affective disorders compared to other psychiatric disorders in children and adolescents with 22q11.2 deletion syndrome. *J Affect Disord*
- Kantrowitz JT, Javitt DC. 2010. N-methyl-d-aspartate (NMDA) receptor dysfunction or dysregulation: The final common pathway on the road to schizophrenia? *Brain Res Bull*
- Karayorgou M, Morris MA, Morrow B, Shprintzen RJ, Goldberg R, et al. 1995. Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. *Proc Natl Acad Sci U S A* 92:7612-6
- Kayadjanian N, Lee HS, Pina-Crespo J, Heinemann SF. 2007. Localization of glutamate receptors to distal dendrites depends on subunit composition and the kinesin motor protein KIF17. *Mol Cell Neurosci* 34:219-30
- Keller MC, Miller G. 2006. Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best? *Behav Brain Sci* 29:385-404; discussion 5-52
- Kety SS, Rosenthal D, Wender PH, Schulsinger F, Jacobsen B. 1976. Mental illness in the biological and adoptive families of adopted individuals who have become schizophrenic. *Behav Genet* 6:219-25

- Kim JS, Kornhuber HH, Schmid-Burgk W, Holzmüller B. 1980. Low cerebrospinal fluid glutamate in schizophrenic patients and a new hypothesis on schizophrenia. *Neurosci Lett* 20:379-82
- Kirov G, Grozeva D, Norton N, Ivanov D, Mantripragada KK, et al. 2009a. Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Hum Mol Genet* 18:1497-503
- Kirov G, Gumus D, Chen W, Norton N, Georgieva L, et al. 2008. Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet* 17:458-65
- Kirov G, Zaharieva I, Georgieva L, Moskvina V, Nikolov I, et al. 2009b. A genome-wide association study in 574 schizophrenia trios using DNA pooling. *Mol Psychiatry* 14:796-803
- Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, et al. 2007. Paired-end mapping reveals extensive structural variation in the human genome. *Science* 318:420-6
- Krebs MO. 2009. [Chromosomal integrity and schizophrenia]. *Med Sci (Paris)* 25:219-21
- Krebs MO, Mouchet S. 2007. [Neurological soft signs and schizophrenia: a review of current knowledge]. *Rev Neurol (Paris)* 163:1157-68
- Kringlen E, Cramer G. 1989. Offspring of monozygotic twins discordant for schizophrenia. *Arch Gen Psychiatry* 46:873-7
- Kristiansen LV, Huerta I, Beneyto M, Meador-Woodruff JH. 2007. NMDA receptors and schizophrenia. *Curr Opin Pharmacol* 7:48-55
- Krystal JH, Anand A, Moghaddam B. 2002. Effects of NMDA receptor antagonists: implications for the pathophysiology of schizophrenia. *Arch Gen Psychiatry* 59:663-4
- Krystal JH, Karper LP, Seibyl JP, Freeman GK, Delaney R, et al. 1994. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry* 51:199-214
- Lau CG, Zukin RS. 2007. NMDA receptor trafficking in synaptic plasticity and neuropsychiatric disorders. *Nat Rev Neurosci* 8:413-26
- Lencz T, Morgan TV, Athanasiou M, Dain B, Reed CR, et al. 2007. Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. *Mol Psychiatry* 12:572-80
- Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, et al. 2003. Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *Am J Hum Genet* 73:34-48
- Lindsay EA, Morris MA, Gos A, Nestadt G, Wolyniec PS, et al. 1995. Schizophrenia and chromosomal deletions within 22q11.2. *Am J Hum Genet* 56:1502-3
- Liu L, Wong TP, Pozza MF, Lingenhoehl K, Wang Y, et al. 2004. Role of NMDA receptor subtypes in governing the direction of hippocampal synaptic plasticity. *Science* 304:1021-4
- Luthi-Carter R, Taylor DM, Pallos J, Lambert E, Amore A, et al. 2010. SIRT2 inhibition achieves neuroprotection by decreasing sterol biosynthesis. *Proc Natl Acad Sci U S A* 107:7927-32
- Lynch M. 2010. Rate, molecular spectrum, and consequences of human mutation. *Proc Natl Acad Sci U S A* 107:961-8

- Mah S, Nelson MR, Delisi LE, Reneland RH, Markward N, et al. 2006. Identification of the semaphorin receptor PLXNA2 as a candidate for susceptibility to schizophrenia. *Mol Psychiatry* 11:471-8
- Malaspina D, Corcoran C, Fahim C, Berman A, Harkavy-Friedman J, et al. 2002. Paternal age and sporadic schizophrenia: evidence for de novo mutations. *Am J Med Genet* 114:299-303
- Manolio TA. 2010. Genomewide association studies and assessment of the risk of disease. *N Engl J Med* 363:166-76
- Marenco S, Weinberger DR. 2000. The neurodevelopmental hypothesis of schizophrenia: following a trail of evidence from cradle to grave. *Dev Psychopathol* 12:501-27
- Margari F, Petruzzelli MG, Lecce PA, Todarello O, De Giacomo A, et al. 2011. Familial liability, obstetric complications and childhood development abnormalities in early onset schizophrenia: a case control study. *BMC Psychiatry* 11:60
- Margari F, Presicci A, Petruzzelli MG, Ventura P, Di Cuonzo F, et al. 2008. Very early onset and greater vulnerability in schizophrenia: A clinical and neuroimaging study. *Neuropsychiatr Dis Treat* 4:825-30
- McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, et al. 2008. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* 40:1166-74
- McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, et al. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 9:356-69
- Meador-Woodruff JH, Healy DJ. 2000. Glutamate receptor expression in schizophrenic brain. *Brain Res Brain Res Rev* 31:288-94
- Mefford HC, Sharp AJ, Baker C, Itsara A, Jiang Z, et al. 2008. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N Engl J Med* 359:1685-99
- Mei L, Xiong WC. 2008. Neuregulin 1 in neural development, synaptic plasticity and schizophrenia. *Nat Rev Neurosci* 9:437-52
- Mill J, Tang T, Kaminsky Z, Khare T, Yazdanpanah S, et al. 2008. Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am J Hum Genet* 82:696-711
- Millar JK, Wilson-Annan JC, Anderson S, Christie S, Taylor MS, et al. 2000. Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum Mol Genet* 9:1415-23
- Mohn AR, Gainetdinov RR, Caron MG, Koller BH. 1999. Mice with reduced NMDA receptor expression display behaviors related to schizophrenia. *Cell* 98:427-36
- Monaghan DT, Jane DE. 2009. Pharmacology of NMDA Receptors.
- Mori H, Manabe T, Watanabe M, Satoh Y, Suzuki N, et al. 1998. Role of the carboxy-terminal region of the GluR epsilon2 subunit in synaptic localization of the NMDA receptor channel. *Neuron* 21:571-80
- Murphy KC, Jones LA, Owen MJ. 1999. High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch Gen Psychiatry* 56:940-5
- Need AC, Ge D, Weale ME, Maia J, Feng S, et al. 2009. A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet* 5:e1000373

- Ng MY, Levinson DF, Faraone SV, Suarez BK, DeLisi LE, et al. 2009a. Meta-analysis of 32 genome-wide linkage studies of schizophrenia. *Mol Psychiatry* 14:774-85
- Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, et al. 2009b. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 461:272-6
- Niemann S, Landers JE, Churchill MJ, Hosler B, Sapp P, et al. 2008. Motoneuron-specific NR3B gene: no association with ALS and evidence for a common null allele. *Neurology* 70:666-76
- Niklasson L, Rasmussen P, Oskarsdottir S, Gillberg C. 2009. Autism, ADHD, mental retardation and behavior problems in 100 individuals with 22q11 deletion syndrome. *Res Dev Disabil* 30:763-73
- Nowak L, Bregestovski P, Ascher P, Herbet A, Prochiantz A. 1984. Magnesium gates glutamate-activated channels in mouse central neurones. *Nature* 307:462-5
- O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, et al. 2008. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* 40:1053-5
- O'Donovan MC, Craddock NJ, Owen MJ. 2009. Genetics of psychosis; insights from views across the genome. *Hum Genet* 126:3-12
- O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, et al. 2011. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet*
- Occhipinti R, Somersalo E, Calvetti D. 2010. Energetics of inhibition: insights with a computational model of the human GABAergic neuron-astrocyte cellular complex. *J Cereb Blood Flow Metab* 30:1834-46
- Parker JA, Arango M, Abderrahmane S, Lambert E, Tourette C, et al. 2005. Resveratrol rescues mutant polyglutamine cytotoxicity in nematode and mammalian neurons. *Nat Genet* 37:349-50
- Pilowsky LS, Bressan RA, Stone JM, Erlandsson K, Mulligan RS, et al. 2006. First in vivo evidence of an NMDA receptor deficit in medication-free schizophrenic patients. *Mol Psychiatry* 11:118-9
- Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, et al. 2010. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 466:368-72
- Pitcher GM, Kalia LV, Ng D, Goodfellow NM, Yee KT, et al. 2011. Schizophrenia susceptibility pathway neuregulin 1-ErbB4 suppresses Src upregulation of NMDA receptors. *Nat Med*
- Porteous DJ, Thomson P, Brandon NJ, Millar JK. 2006. The genetics and biology of DISC1--an emerging role in psychosis and cognition. *Biol Psychiatry* 60:123-31
- Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, et al. 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460:748-52
- Rapoport JL, Addington AM, Frangou S, Psych MR. 2005. The neurodevelopmental model of schizophrenia: update 2005. *Mol Psychiatry* 10:434-49
- Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, et al. 2006. Global variation in copy number in the human genome. *Nature* 444:444-54
- Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, et al. 2010. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* 328:636-9

- Ross CA, Margolis RL, Reading SA, Pletnikov M, Coyle JT. 2006. Neurobiology of schizophrenia. *Neuron* 52:139-53
- Rujescu D, Ingason A, Cichon S, Pietilainen OP, Barnes MR, et al. 2009. Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum Mol Genet* 18:988-96
- Salter MW, Dong Y, Kalia LV, Liu XJ, Pitcher G. 2009. Regulation of NMDA Receptors by Kinases and Phosphatases.
- Sebat J. 2007. Major changes in our DNA lead to major changes in our thinking. *Nat Genet* 39:S3-5
- Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, et al. 2007. Strong association of de novo copy number mutations with autism. *Science* 316:445-9
- Sebat J, Lakshmi B, Troge J, Alexander J, Young J, et al. 2004. Large-scale copy number polymorphism in the human genome. *Science* 305:525-8
- Setou M, Nakagawa T, Seog DH, Hirokawa N. 2000. Kinesin superfamily motor protein KIF17 and mLin-10 in NMDA receptor-containing vesicle transport. *Science* 288:1796-802
- Shastri BS. 1999. Recent developments in the genetics of schizophrenia. *Neurogenetics* 2:149-54
- Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, et al. 2009. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 460:753-7
- Shianna KV, Willard HF. 2006. Human genomics: in search of normality. *Nature* 444:428-9
- Shifman S, Johannesson M, Bronstein M, Chen SX, Collier DA, et al. 2008. Genome-wide association identifies a common variant in the reelin gene that increases the risk of schizophrenia only in women. *PLoS Genet* 4:e28
- Sklar P, Pato MT, Kirby A, Petryshen TL, Medeiros H, et al. 2004. Genome-wide scan in Portuguese Island families identifies 5q31-5q35 as a susceptibility locus for schizophrenia and psychosis. *Mol Psychiatry* 9:213-8
- St Clair D, Blackwood D, Muir W, Carothers A, Walker M, et al. 1990. Association within a family of a balanced autosomal translocation with major mental illness. *Lancet* 336:13-6
- Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, et al. 2009. Common variants conferring risk of schizophrenia. *Nature* 460:744-7
- Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, et al. 2008. Large recurrent microdeletions associated with schizophrenia. *Nature* 455:232-6
- Sullivan PF, Kendler KS, Neale MC. 2003. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry* 60:1187-92
- Sullivan PF, Lin D, Tzeng JY, van den Oord E, Perkins D, et al. 2008. Genomewide association for schizophrenia in the CATIE study: results of stage 1. *Mol Psychiatry* 13:570-84
- Sun J, Jia P, Fanous AH, van den Oord E, Chen X, et al. 2010. Schizophrenia gene networks and pathways and their applications for novel candidate gene selection. *PLoS One* 5:e11351
- Svensson AC, Lichtenstein P, Sandin S, Hultman CM. 2007. Fertility of first-degree relatives of patients with schizophrenia: a three generation perspective. *Schizophr Res* 91:238-45

- Tarabeux J. 2010. De Novo Truncating Mutation in Kinesin 17 Associated with Schizophrenia. *Biol Psychiatry*
- Tuominen HJ, Tiihonen J, Wahlbeck K. 2006. Glutamatergic drugs for schizophrenia. *Cochrane Database Syst Rev*:CD003730
- Van Snellenberg JX, de Candia T. 2009. Meta-analytic evidence for familial coaggregation of schizophrenia and bipolar disorder. *Arch Gen Psychiatry* 66:748-55
- Vincent A, Bien CG. 2008. Anti-NMDA-receptor encephalitis: a cause of psychiatric, seizure, and movement disorders in young adults. *Lancet Neurol* 7:1074-5
- Visser LE, de Ligt J, Gilissen C, Janssen I, Stehouwer M, et al. 2010. A de novo paradigm for mental retardation. *Nat Genet* 42:1109-12
- Voglis G, Tavernarakis N. 2006. The role of synaptic ion channels in synaptic plasticity. *EMBO Rep* 7:1104-10
- Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, et al. 2008. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 320:539-43
- Watis L, Chen SH, Chua HC, Chong SA, Sim K. 2008. Glutamatergic abnormalities of the thalamus in schizophrenia: a systematic review. *J Neural Transm* 115:493-511
- Weinberger DR. 1987. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 44:660-9
- Weinberger DR, McClure RK. 2002. Neurotoxicity, neuroplasticity, and magnetic resonance imaging morphometry: what is happening in the schizophrenic brain? *Arch Gen Psychiatry* 59:553-8
- Wender PH, Rosenthal D, Kety SS, Schulsinger F, Welner J. 1974. Crossfostering. A research strategy for clarifying the role of genetic and experiential factors in the etiology of schizophrenia. *Arch Gen Psychiatry* 30:121-8
- Williams HJ, Craddock N, Russo G, Hamshere ML, Moskvina V, et al. 2011. Most genome-wide significant susceptibility loci for schizophrenia and bipolar disorder reported to date cross-traditional diagnostic boundaries. *Hum Mol Genet* 20:387-91
- Williams HJ, Owen MJ, O'Donovan MC. 2007. Is COMT a susceptibility gene for schizophrenia? *Schizophr Bull* 33:635-41
- Williams K. 2009. Extracellular Modulation of NMDA Receptors.
- Williams NM, Bowen T, Spurlock G, Norton N, Williams HJ, et al. 2002. Determination of the genomic structure and mutation screening in schizophrenic individuals for five subunits of the N-methyl-D-aspartate glutamate receptor. *Mol Psychiatry* 7:508-14
- Wong RW, Setou M, Teng J, Takei Y, Hirokawa N. 2002. Overexpression of motor protein KIF17 enhances spatial and working memory in transgenic mice. *Proc Natl Acad Sci U S A* 99:14500-5
- Xi ZX, Baker DA, Shen H, Carson DS, Kalivas PW. 2002. Group II metabotropic glutamate receptors modulate extracellular glutamate in the nucleus accumbens. *J Pharmacol Exp Ther* 300:162-71
- Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. 2008. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet* 40:880-5

- Xu B, Woodroffe A, Rodriguez-Murillo L, Roos JL, van Rensburg EJ, et al. 2009. Elucidating the genetic architecture of familial schizophrenia using rare copy number variant and linkage scans. *Proc Natl Acad Sci U S A* 106:16746-51
- Yang J, Visscher PM, Wray NR. 2010. Sporadic cases are the norm for complex disease. *Eur J Hum Genet* 18:1039-43
- Yin X, Takei Y, Kido MA, Hirokawa N. 2011. Molecular Motor KIF17 Is Fundamental for Memory and Learning via Differential Support of Synaptic NR2A/2B Levels. *Neuron* 70:310-25

Etude des déterminants génétiques des psychoses à début précoce

Génétique de la schizophrénie et hypothèse glutamatergique

Les troubles schizophréniques (SCZ) ont une forte héritabilité, de l'ordre de 80%, mais, une très faible part du risque génétique a été identifiée. La plupart des études ont considéré l'implication de polymorphismes fréquents, chacun ayant un effet relativement faible individuellement, alors que les études de variants du nombre de copies (CNVs) ainsi que les études d'anomalies chromosomiques ont pointé l'implication possible de variants rares et *de novo* à une forte pénétrance.

Dans une première partie, nous présentons une synthèse sur les facteurs génétiques dans la SCZ, puis une revue des arguments en faveur de l'implication d'anomalies du système glutamatergique dans la SCZ, domaine sur lequel s'est centré notre travail.

Notre travail s'inscrit dans un projet plus vaste, Synapse to Disease (S2D) ayant pour objectif de séquencer 1000 gènes synaptiques dans des cohortes de patients atteints de schizophrénie ou de troubles du spectre autistique. Nous avons exploré en particulier le système glutamatergique et les récepteurs NMDA. Dans un premier article, nous montrons une association d'une mutation tronquante *de novo* de la kinésine 17, impliquée dans le transport de la sous-unité GRIN2B des récepteurs NMDA. Dans un second article, nous explorons les mutations rares et *de novo* dans les sous-unités des récepteurs NMDA et montrons l'association de mutation *de novo* dans *GRIN2A* et *GRIN2B* avec des cas de SCZ et d'autisme. Nos résultats renforcent l'idée qu'une part des cas de schizophrénie pourrait être due à l'implication de mutations rare à effet majeur, hypothèse alternative mais non exclusive à l'hypothèse d'interactions entre variants génétiques fréquents à effet mineur.

Study of genetic factors in early onset psychosis.

Genetic of schizophrenia and glutamatergic hypothesis.

Schizophrenic disorders (SCZ) have high heritability (around 80%), but only a small part has been characterized. Most studies have focussed on common polymorphisms, each having small individual effect, whereas copy number variant and chromosomal abnormalities studies have pointed to the possible involvement of rare and *de novo* mutations with high penetrance.

In the first part of this manuscript, we will present a synthesis on genetic factors of SCZ and then a review of the arguments supporting an involvement of glutamatergic system abnormalities in SCZ, which is the focus of our research.

Our work is part of a global project, Synapse to Disease (S2D), that aimed to sequence 1000 synaptic genes in cohort of patients affected with schizophrenia or autism spectrum disorders. We focussed in particular on the glutamatergic system and NMDA receptors. In a first publication we show an association between SCZ and a *de novo* truncating mutation of kinesin 17, which has been implicated in the transport of the GRIN2B subunit of NMDA receptors. In a second publication we explore rare and *de novo* mutations in NMDA receptor subunits. We show an association between *de novo* mutations in *GRIN2A* and *GRIN2B* with cases of SCZ and autism. Our results strengthen the idea that a portion of schizophrenia cases could be related to rare mutations having a high penetrance, an alternative but not contradictory explanation to the hypothesis for an interaction between common variants having a small effect.